1-1-2015

Essays on Bayesian Macroeconometrics

Molin Zhong
University of Pennsylvania, molin@sas.upenn.edu

Essays on Bayesian Macroeconometrics

Abstract
This dissertation consists of three chapters that study the determinants of macroeconomic fluctuations, with a particular emphasis on the roles of agents' expectations and assessments of risks. In the first chapter, I study a business cycle model where the probability of transitioning to a downturn state characterized by low growth evolves over time. I call a change in the future probability of transitions to the downturn state a downturn risk shock. An increase in the risk of the downturn state leads to declines in consumption, investment, output, and hours. I take the model to the data using Bayesian methods. The fluctuations caused by expectations changes from the downturn risk shock account for substantial output variations at business cycle frequencies and hours fluctuations at medium run frequencies. The extracted time-varying probability process from the model matches well with the University of Michigan Index of Consumer Sentiments (ICS). Impulse response functions from downturn risk shocks produce the same comovement as those from ICS innovations in a structural vector autoregression. The second chapter, co-authored with Minchul Shin, proposes the multivariate stochastic volatility in vector autoregression model as a framework for studying the real effects of uncertainty shocks. We advance a new approach to structurally identify uncertainty shocks that does not rely on identification of the level structural shocks. We estimate the model using a Bayesian Markov chain Monte Carlo algorithm, and we show how to construct impulse response functions and variance decompositions that can be used to analyze identified uncertainty shocks. The third chapter, also co-authored with Minchul Shin, suggests using ``realized volatility'' as a volatility proxy to aid in model-based multivariate bond yield density forecasting. To do so, we develop a general estimation approach to incorporate volatility proxy information into dynamic factor models with stochastic volatility. We study the density prediction performance on U.S. bond yields of including realized volatility into a dynamic Nelson-Siegel (DNS) model with stochastic volatility. The results clearly indicate that using realized volatility improves density forecasts relative to popular specifications in the DNS literature that neglect realized volatility.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
Jesus Fernandez-Villaverde

Keywords
Bayesian methods, Business cycles, Dynamic factor model, Forecasting, Stochastic volatility, Vector autoregression

ESSAYS IN BAYESIAN MACROECONOMETRICS

Molin Zhong

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2015

Supervisor of Dissertation

Jesús Fernández-Villaverde
Professor of Economics

Graduate Group Chairperson

George Mailath, Professor of Economics

Dissertation Committee

Jesús Fernández-Villaverde, Professor of Economics
Francis X. Diebold, Professor of Economics
Frank Schorfheide, Professor of Economics

ESSAYS IN BAYESIAN MACROECONOMETRICS

To my family

ACKNOWLEDGEMENTS

I cannot express how grateful I am to have Professor Jesús Fernández-Villaverde as my advisor. Professor Fernández-Villaverde guided me through the arduous process of becoming a researcher in economics. He always was there to offer me his advice and wisdom. Most importantly, through my time with him, I have learned the importance of positive thinking and belief in myself. These are lessons that I will take with me beyond the profession and I thank him dearly for instilling them in me.

I would like to thank Professors Francis Diebold and Frank Schorfheide. Professor Diebold spent the time to introduce me into the profession when I was still an undergraduate. He provided me the opportunity to attend Penn and I could always count on him to give helpful guidance and productive comments. Professor Schorfheide pushed me further to improve my work. He provided invaluable feedback on my dissertation to help me take it to the level it is at today. I am lucky to have them on my committee.

My gratitude goes out to the members of the Penn Econometrics lunch group and the economists at the Federal Reserve Banks of Philadelphia and St. Louis. I thank both the Federal Reserve Bank of Philadelphia and the Federal Reserve Bank of St. Louis for their hospitality.

I am especially indebted to Pablo Guerron-Quintana, who provided me useful direction and support throughout my final years in the program. I would also like to thank George Alessandria, Xu Cheng, Frank DiTraglia, Thorsten Drautzburg, and Edison Yu for many helpful comments and feedback.

Thank you to the wonderful friends I have made along the process: Alessandro Arloto, Lorenzo Braccini, Yumi Koh, Laura Liu, Yang Liu, Daniel Neuhann, Qiusha Peng, Devin Reilly, and Minchul Shin. I would especially like to thank Lorenzo for being a wonderful

roommate for three years and Devin for being a great officemate for four years. I am exceptionally thankful to my coauthor Minchul. Quoting the words of my advisor about his coauthor: you are "the best coauthor I could have hoped for."

Finally, I would like to thank my parents and sister for their support during my time at Penn. I could not have made it this far without them.

Molin Zhong

Philadelphia, PA

April 27, 2015

ABSTRACT

ESSAYS IN BAYESIAN MACROECONOMETRICS

Molin Zhong

Jesús Fernández-Villaverde

This dissertation consists of three chapters that study the determinants of macroeconomic fluctuations, with a particular emphasis on the roles of agents' expectations and assessments of risks. In the first chapter, I study a business cycle model where the probability of transitioning to a downturn state characterized by low growth evolves over time. I call a change in the future probability of transitions to the downturn state a downturn risk shock. An increase in the risk of the downturn state leads to declines in consumption, investment, output, and hours. I take the model to the data using Bayesian methods. The fluctuations caused by expectations changes from the downturn risk shock account for substantial output variations at business cycle frequencies and hours fluctuations at medium run frequencies. The extracted time-varying probability process from the model matches well with the University of Michigan Index of Consumer Sentiments (ICS). Impulse response functions from downturn risk shocks produce the same comovement as those from ICS innovations in a structural vector autoregression. The second chapter, co-authored with Minchul Shin, proposes the multivariate stochastic volatility in vector autoregression model as a framework for studying the real effects of uncertainty shocks. We advance a new approach to structurally identify uncertainty shocks that does not rely on identification of the level structural shocks. We estimate the model using a Bayesian Markov chain Monte Carlo algorithm, and we show how to construct impulse response functions and variance decompositions that can be used to analyze identified uncertainty shocks. The third chapter, also co-authored with Minchul Shin, suggests using "realized volatility" as a volatility proxy to aid in model-based multivariate bond yield density forecasting. To do so, we develop a general estimation approach to incorporate volatility proxy information into dynamic

factor models with stochastic volatility. We study the density prediction performance on U.S. bond yields of including realized volatility into a dynamic Nelson-Siegel (DNS) model with stochastic volatility. The results clearly indicate that using realized volatility improves density forecasts relative to popular specifications in the DNS literature that neglect realized volatility.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1 : Variable Downturn Risk

1.1. Introduction

In this paper, I study a business cycle model with variable downturn risk. This study is motivated by the asymmetric nature of macroeconomic fluctuations. On U.S. data, output, consumption, and investment growth decline rapidly during recessions but increase gradually during expansions. Moreover, the chance of a recession changes over time. The economy goes through decade-long stretches with weak growth and frequent, deep recessions and decade-long stretches with strong growth and infrequent, mild recessions. Do agents react to this variable risk of downturns? How important are agents' assessments of downturn risk to macroeconomic fluctuations?

I address these questions as follows. I first provide stylized facts suggestive of the presence of downturn risk. Current dynamic equilibrium models have a difficult time accounting for this new form of risk. I therefore add a tightly-parametrized time-varying probability of transitions to a downturn state over which agents have rational expectations to an otherwise standard business cycle model.

A variety of stylized facts point to the presence and importance of downturn risk. Distributions of output, consumption, and investment growth show a long downside tail, indicating a downturn state with rapid declines in the variables. These rapid declines do not seem to be dispersed uniformly across time but instead appear to have a component operative at longer horizons associated with the medium-frequency cycle ($32 - 200$ quarters). For instance, the 1970s and early 1980s, a time period of high risk of energy shortages, had four deep recessions with sharp decreases in the macroeconomic aggregates. On the other hand, during the late 1980s into the mid 2000s, a time period that included the information technology revolution, the U.S. only experienced two mild recessions with no sharp declines in output.

I also present evidence that agents react to this downturn risk. I show that the University of Michigan Index of Consumer Sentiments (ICS), a survey-based index of consumer expectations, forecasts the lower tail of one-quarter ahead real per capita consumption growth innovations but not the median or upper tail. This asymmetry suggests that information relevant for downturns in the economy, more so than for upturns, shape agents' expectations. Moreover, fluctuating agents' expectations and beliefs themselves seem to matter. In a vector autoregression (VAR) with the ICS and a variety of key business cycle variables, identified innovations to the ICS lead to significant movements in macroeconomic variables. Along with documenting the importance of agents' expectations for business cycle movements, the VAR also gives suggestive evidence as to the comovements expectations shocks should produce in a structural model.

The model I present allows for two states of the world: a downturn state and a normal state. I introduce a tightly-parametrized first-order Markov time-varying probability of transitioning to the downturn state that agents understand. This time-varying transition probability captures the changing downturn risks that agents face. Moreover, a highly persistent probability process implies that the shock has implications for both business cycle and medium-frequency cycle fluctuations. The modeling framework is flexible enough to allow for downturn, upturn, or symmetric risks to operate, but the econometric estimation clearly points to variable downturn risks as driving business cycles. The downturn state in the model has low productivity growth and high depreciation whereas the normal state has high productivity growth and low depreciation. A possible interpretation of low productivity growth during downturns is that capital misallocation increases. High depreciation during downturns could result from increased economic depreciation or obsolescence, as opposed to physical depreciation, of the capital stock.

A shock to the probability of the downturn state occurring, which I call a *downturn risk shock*, falls into the category of expectations shocks. Expectations shocks emphasize changing agents' expectations in driving macroeconomic fluctuations with no corresponding move-

ments in observed economic fundamentals (for example: productivity, preferences, or policy). In line with the literature, my VAR results suggest that in response to an identified negative innovation to the ICS, both consumption and investment decline (see, for example, Beaudry and Portier (2006)). While expectations shocks in general have a notoriously difficult time producing comovement between consumption and investment in standard business cycle models, as shown by Beaudry and Portier (2004), the *pure downturn risk shock*, which captures a change in risk of the downturn state with no subsequent movements in productivity or depreciation (a *pure* change in expectations), can generate such comovement in an otherwise standard real business cycle model. Expected future productivity growth declines reduce the incentive to consume immediately through the wealth effect but have difficulties inducing a decline in investment. Expected future depreciation increases reduce the incentive to invest, which help produce comovement between output, consumption, and investment. Using the same model with the same parameterization, I show a variety of news shocks cannot generate this comovement, illustrating the importance of having both productivity growth and depreciation in the downturn risk shock.

Since my model falls into the class of Markov-switching dynamic equilibrium models, I can solve it using the perturbation methodology as in Foerster, Rubio-Ramirez, Waggoner, and Zha (2013). The time-varying probability process governing the downturn risk is a discretization of an underlying continuous autoregressive process parametrized by a persistence, standard deviation, and long-run mean parameter. In this way, I can account for rich dynamics in downturn risk while keeping the model tightly parametrized.

I take a quantitative version of the model with competing structural shocks to U.S. macroeconomic data from $1960Q2 - 2011Q4$ using Bayesian methods.

In the estimated model, the downturn risk shock is an important driver of fluctuations at both the business and medium-frequency cycles. At business cycle frequencies ($8 - 32$ quarters), the downturn risk shock plays a dominant role in consumption fluctuations, accounting for around 70 percent of fluctuations, while being an important determinant of

output (around 30 percent of fluctuations) and hours (around 25 percent of fluctuations) movements. The downturn risk in the economy triggers a wealth effect that becomes a main driving force of consumption. Moving to medium frequencies ($32 - 200$ quarters), as defined by Comin and Gertler (2006), the downturn risk shock becomes the dominant driver of the macroeconomic aggregates, accounting for around 60 percent of output movements. A downturn probability shock can be seen as a type of intertemporally correlated news shock. Its propagation has two sources of persistence: persistence in the news and persistence in the realizations. Therefore, the effects of a change in downturn risk can persist for a long time, which make the shock a natural candidate for driving longer-horizon fluctuations. Periods of frequent and deep recessions have high downturn risk with many transitions to the downturn state. Because agents also understand the high downturn risk present in the economy, the lower expectations produce a further drag on macroeconomic performance. On the other hand, when the economy has low downturn risk, there are few transitions to the downturn state, generating infrequent recessions. Agents also understand this low risk, which further boosts economic activity.

The pure downturn risk shock is empirically important for business cycles and drives hours fluctuations at lower frequencies. Moreover, an increase in the probability of the downturn state associated with the pure downturn risk shock leads to transitory declines in output, consumption, investment, and hours, a hallmark of business cycle fluctuations. The pure downturn risk shock drives around 40 percent of the movements in hours at medium frequencies, which drastically decreases the importance of the labor supply shock in matching hours. The model therefore gives a new interpretation to labor supply fluctuations: they come about due to changes in the assessments of agents of the downturn risk in the economy. Evidence outside of the structural model seems to suggest such a theory of hours fluctuations. The ICS and hours comove well across the data sample. Moreover, variance decomposition results from a structural VAR suggest that up to 70% of hours fluctuations 40-quarters out originate from identified ICS innovations.

Finally, the estimated model aligns well with exogenous supporting evidence. The smoothed downturn risk probability process comoves well with the Index of Consumer Sentiments. In addition, the impulse response to the downturn risk shock generates the same comovement in macroeconomic variables as the response to an innovation in the ICS from a structurally-identified VAR, which is further reassuring for the interpretation. A competing productivity growth news shock in the model, for instance, misses the investment and federal funds rate movements. Moreover, a variety of estimated models with alternative popular specifications for expectations shocks also do not match these comovements.

Given the importance of the downturn risk shock in historical business cycle fluctuations, I then evaluate how well it can account for the recent Great Recession. This recession is interesting because it represents a large break in the economy. The previous 20 years were characterized by the Great Moderation, a time of lower variation in macroeconomic series and two mild recessions. Many economists think of the Great Recession as marking the end of the Great Moderation. Therefore, changing expectations and sentiments of agents could play a large role in this uncertain time period. The downturn risk shock is an important driver of movements in output, consumption, and hours across the recession. Shutting down all of the other shocks and only having downturn risks operative leads to movements in the three series consistent with the data. The pure downturn risk shock accounts for between $25 - 40\%$ of the decline of the three series across the recession.

In summary, changes in agents' assessments of the downturn risk in the economy drive macroeconomic fluctuations in my model. I find that these downturn risks have an important component operative at medium-frequency horizons while having significant implications at business cycle frequencies. Moreover, changes in expectations alone caused by the evolving risks are empirically relevant for business and medium-frequency cycle macroeconomic fluctuations. These expectations changes are important because they produce comovement between consumption, investment, output, and hours, a defining feature of business cycle fluctuations.

Related literature My work connects to the literature in five main areas.

First, there exists a large literature discussing the importance of signals about the future for business cycle fluctuations. This literature can be divided into three general categories: news, noise, and sentiments.

The news shocks of Beaudry and Portier (2004) capture the idea that agents respond to future fundamentals in making their decisions. Davis (2007), Schmitt-Grohe and Uribe (2012), Fujiwara, Hirose, and Shintani (2011), and Khan and Tsoukalas (2012) estimate models evaluating the importance of news shocks in driving business cycles. The downturn risk shock can be considered a type of news shock. It differs from the standard news shock in four key aspects. First, the process underlying the downturn risk shock in principle allows for downside, upside, or symmetric risks, whereas the news shock only considers the symmetric case. Second, the papers above model news as certain events in the future. The downturn risk shock, on the other hand, only changes the probabilities with which future states of the world will occur. This allows the downturn risk shock to also account for situations where the agents' expectations change but the realized future observed fundamentals do not. Third, the downturn risk shock combines two different types of news: productivity growth and depreciation. I show this combination is key to generate comovement between consumption and investment without resorting to large adjustment costs. Fourth, the downturn risk shock is a parsimonious way of modeling intertemporal correlation of news shocks, as will be discussed in more detail.

Noise shocks of Barsky and Sims (2012) and Blanchard, L'Huillier, and Lorenzoni (2013) characterize a situation where agents receive noisy signals about the future. A noise shock changes agents' expectations while not impacting the observed fundamentals of the economy as the noise only affects the signals that agents receive. A pure downturn risk shock has similarities to a noise shock in that the observed fundamentals of the economy do not change but the agents' expectations do. Along with capturing asymmetric risks, the pure downturn risk shock has longer-horizon implications for fluctuations than what is typically empirically

found for a noise shock.

Angeletos and La'O (2013) develop a theory of sentiment-driven fluctuations where dispersed information opens the door to changes in agents' beliefs as a driver of macroeconomic aggregates. Angeletos, Collard, and Dellas (2014) embed this beliefs shock into real business cycle and New Keynesian models and find that it is an important driver of fluctuations. The downturn risk shock, in contrast to the sentiments shock, changes agents' expectations through the time-varying risks of a bad state of the world occurring. It does not require dispersed information.

Second, my work connects to the disasters literature introduced into macroeconomic models by Gourio (2012). In a theoretical model, he shows that time-varying probability of disasters can explain a variety of asset price phenomena while simultaneously driving business cycle movements. As opposed to the disaster states that Gourio (2012) considers, my model focuses on time-varying risks of less extreme downturn states that occur during recessions. The increased frequency of these downturn states observed in the data relative to the frequency of disaster states facilitates econometric estimation.

Third, the downturn risk shock links to the idea of long run risks of Bansal and Yaron (2004) and medium-frequency cycles as in Comin and Gertler (2006). In the context of empirical dynamic equilibrium models, these ideas are usually modeled as trend productivity growth shocks, as in Croce (2014), Barsky and Sims (2012), and Blanchard, L'Huillier, and Lorenzoni (2013). Relative to these models, my model ties the medium-frequency cycle to business cycle asymmetries. Also, the empirically relevant inclusion of a depreciation component in the downturn risk shock allows for comovement between consumption and investment from pure changes in expectations.

Fourth, on the methodological side, this paper relates to the literature on estimating Markov-switching dynamic equilibrium models (Schorfheide (2005), Liu, Waggoner, and Zha (2011), Bianchi (2013), and Foerster, Rubio-Ramirez, Waggoner, and Zha (2013)).

That literature builds on the work of Hamilton (1989), which first advances the Markov-switching model. Diebold, Lee, and Weinbach (1994) propose the Markov-switching model with time-varying transition probabilities, which has gained much importance in reduced-form work. Most papers in the area specify an observable variable driving the transition probabilities (Filardo (1994)). By embedding the Markov-switching time-varying transition probability process into a dynamic equilibrium model, it is possible to use the decisions of the agents to infer the transition probabilities at a given point in time. Bianchi and Melosi (2013) discuss how the Markov-switching framework can quite naturally model movements in agents' expectations but does not conduct any empirical evaluation. My paper shows the empirical relevance of using Markov-switching methods to model agents' expectations.

Finally, this paper relates to the literature on generating asymmetry in business cycle models. These papers fall into two broad classes: ones that generate the asymmetry using exogenous shocks and ones that do so endogenously in the model. As my paper does not focus on generating asymmetry in macroeconomic models, but rather on producing expectations-driven business cycles via the interaction of time-varying downturn risks and frictions in the model, it falls into the former class. Andreasen, Fernandez-Villaverde, and Rubio-Ramirez (2013), Ruge-Murcia (2012a), and Ruge-Murcia (2012b) use innovations drawn from a skewed distribution to capture the skewness in macroeconomic and financial data. Relative to that literature, my framework can also account for the predictability of downside shocks. It can capture the notion of time-varying risks of downside shocks and agents reacting to these changing risks. Considering effects of asymmetric adjustment costs (Kim and Ruge-Murcia (2009) and Aruoba, Bocola, and Schorfheide (2013a)), the zero lower bound on macroeconomic fluctuations (Fernandez-Villaverde, Gordon, Guerron-Quintana, and Rubio-Ramirez (2012) and Aruoba, Cuba-Borda, and Schorfheide (2013)), and the impact of occasionally binding financial constraints (Guerrieri and Iacoviello (2013)), are examples of models that generate macroeconomic asymmetries endogenously.

The plan of the paper is as follows. Section 2 discusses the stylized facts. Section 3 illustrates

the mechanism by embedding it within a real business cycle model. Section 4 discusses the solution method and conducts Bayesian estimation on a quantitative dynamic equilibrium model. Finally, section 5 concludes.

1.2. Stylized facts

In this section, I review four stylized facts about business cycle fluctuations that are relevant to the topic at hand. These stylized facts provide evidence that (1) a downturn state into which the economy occasionally transitions exists, (2) the risks of transitions to the downturn state change over time, (3) these downturn risks impact agents' sentiments and expectations, and (4) changes in agents' sentiments and expectations in and of themselves are an important driver of macroeconomic fluctuations.

1. Output, consumption, and investment growth are negative skewed.

   One defining feature of the U.S. business cycle is the asymmetry across expansions and recessions. Expansions tend to produce gradual increases in growth rates of macroeconomic aggregates whereas recessions lead to sharp decreases. Many researchers have noted this asymmetry in reduced-from work, including Neftci (1984), Hamilton (1989), Morley and Piger (2012), and Aruoba, Bocola, and Schorfheide (2013a). From an unconditional perspective, the asymmetric business cycle leads to the aggregate series of output, consumption, and investment growth exhibiting clear negative skewness[1].

Series	Skewness	Kurtosis	J-B test p-val
Output growth	-0.64	4.70	< 0.001
Consumption growth	-0.95	6.45	< 0.001
Investment growth	-0.41	4.43	< 0.001

Table 1 Quarterly per capita output, consumption, and investment growth $1953Q2-2011Q4$

[1] Details on the construction of the data can be found in the appendix.

Table 1 displays the skewness and kurtosis of these three macroeconomic aggregates along with Jarque-Bera tests of normality. They all feature strong negative skewness and fail normality tests.

Figure 1 shows the forecast errors from fitting a simple AR(1) model to output growth from $1953Q2 - 2011Q4$. The forecast errors display a long left tail, mirroring the skewness found in the output growth data itself. Zooming in on this left tail, there are 7 instances where the forecast errors fall less than 2.5 standard deviations below the mean. This empirical fact suggests the presence of a downturn state into which the economy occasionally transitions. For example, a normal distribution would rarely generate so many observations so far below the mean[2].

2. Frequencies and depths of recessions change over time.

The first stylized fact gives an unconditional statement about the asymmetry of the U.S. business cycle. Looking beyond an unconditional perspective reveals the long-horizon, changing frequency of rapid declines in the macroeconomic variables, as pointed out by Comin and Gertler (2006) and dubbed the medium-frequency cycle. They suggest that the economy enters decade-long stretches with poor macroeconomic performance and frequent, deep recessions followed by decade-long stretches of robust macroeconomic performance with infrequent, mild recessions. Across the 1950s into the early 1960s, the U.S. went through three recessions with large declines in output and investment growth. The mid-1960s, on the other hand, had strong economic growth and was recession-free. From $1969 - 1982$, the U.S. again reverted to a time period of poor economic performance. In a 13-year stretch, the U.S. went through four recessions that produced large decreases in output and investment growth as well as a dramatic slowing of consumption growth. For the next 23 years, from $1983 - 2006$, there were only two relatively mild recessions. Finally, the Great Recession was one

[2]In 10000 samples of 235 observations (the sample size) generated from a standard normal distribution, no samples had 7 or more observations below 2.5 standard deviations below the mean.

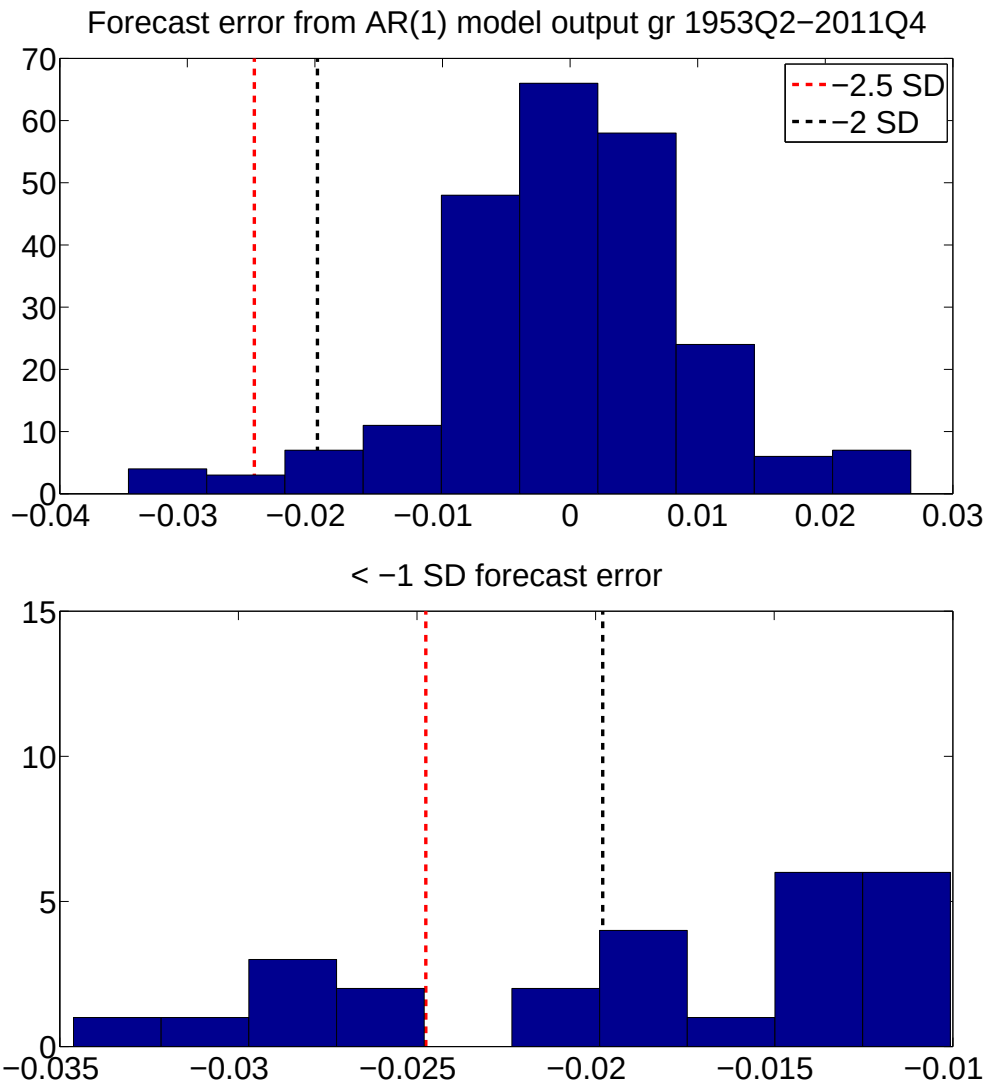


Figure 1 Forecast errors from AR(1) model fit with OLS to output growth. The top panel shows the histogram of the forecast errors across the entire sample period, whereas the bottom panel zooms in on the left tail. The red dashed line gives the 2.5 standard deviations below the mean line while the black dashed line gives the 2 standard deviation line.

of the deepest recessions in the sample, producing declines in the aggregates much larger than the "average" recession during any other period.

Series	$1953-1961$ (3 recessions)	$1961-1969$ (0 recessions)	$1969-1982$ (4 recessions)
Out gr	-0.82%	$--$	-0.87%
Cons gr	0.35%	$--$	0.05%
Inv gr	-3.99%	$--$	-3.53%
Series	$1983-2006$ (2 recessions)	Great Recession	
Out gr	-0.45%	-1.41%	
Cons gr	0.26%	-0.41%	
Inv gr	-2.70%	-5.57%	

Table 2 Average quarterly declines in macroeconomic aggregates across NBER-designated recession periods

Moreover, deep recessions are associated with large downside shocks hitting the economy. Figure 2 shows that of the 7 recessions in the 1950s to early 1960s and late 1960s to early 1980s time frames, all but one are associated with innovations in output growth lower than 2 standard deviations below the mean ($< -2.0 \times 10^{-2}$). The two mild recessions across the 1990s and 2000s, in contrast, do not have any such deep downside innovations.

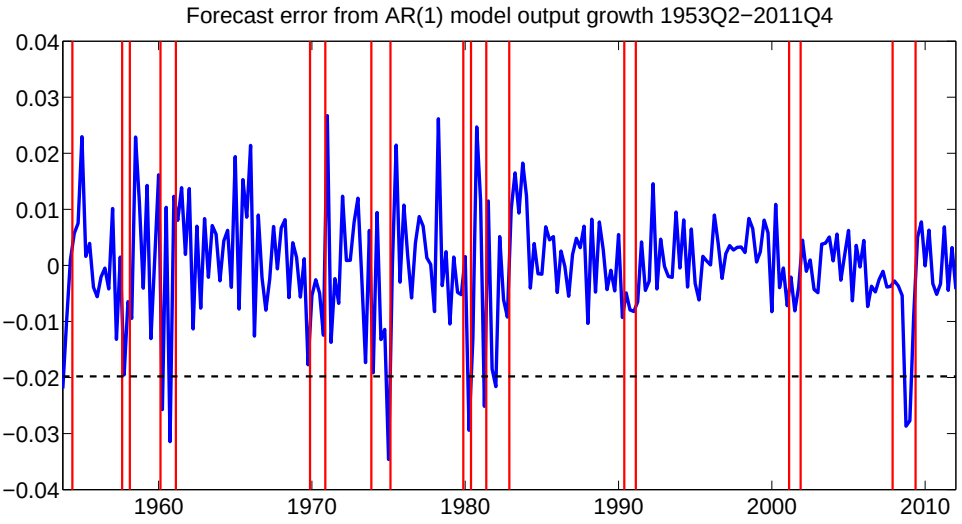


Figure 2 Time series of forecast errors from AR(1) model fit with OLS to output growth. The red vertical lines denote NBER-dated recessions. The black horizontal dashed line gives the 2 standard deviations below the mean line.

3. Consumer sentiments movements are more associated with downturns in future consumption growth innovations than upturns.

   The University of Michigan Index of Consumer Sentiments is a popular proxy for expectations and beliefs of consumers[3]. Barsky and Sims (2012) have shown a relationship between the ICS and news about the future, suggesting that overall, the ICS has predictive content about future fundamentals. In this section, I document an asymmetry of this predictive relationship by illustrating that the ICS forecasts the 20th percentile of consumption growth innovations significantly better than the historical quantile, but it does not forecast the 80th percentile significantly better. This asymmetric relationship between the ICS and future downside innovations suggests that information about downside risk in the economy is more relevant in forming agents' expectations relative to information about upside risk.

   First, I fit an AR(1) with ordinary least squares to per capita consumption growth

[3]Papers such as Barsky and Sims (2012) and Angeletos, Collard, and Dellas (2014) have used elements of the University of Michigan Index of Consumer Sentiments as a proxy for expectations of consumers.

Quantile	Relative Loss
20	0.90**
50	0.96
80	1.00

Table 3 Quantile regression forecast results for time $t+1$ consumption growth innovation given time t log Index of Consumer Sentiments. Using one-sided tests of higher predictive power of the sentiments model, $**$ indicates significance at the 5% level, and $*$ indicates significance at the 10% level.

data and extract the fitted innovations[4]. Then, I perform a quantile regression of time t consumption growth innovations ($\hat{\epsilon}_t^c$) on time $t-1$ log of consumer sentiments. Equation 1.1 gives the quantile regression specification for the τth quantile. My sample runs from $1960Q2 - 2011Q4$. ICS data availability restricts the beginning date of the sample. I begin one-quarter ahead forecasting at $1985Q3$ and do an expanding window recursive estimation until the end of the sample.

$$Q_{\hat{\epsilon}_t^c}\left(\tau | \log sent_{t-1}\right) = \alpha(\tau) + \beta(\tau) \log sent_{t-1} \tag{1.1}$$

$$\frac{1/N \sum_n \left[\rho_\tau \left(\hat{\epsilon}_{t+1}^c - \tilde{\alpha}(\tau) - \tilde{\beta}(\tau) \log sent_t\right)\right]}{1/N \sum_n \left[\rho_\tau \left(\hat{\epsilon}_{t+1}^c - \tilde{q}(\tau)\right)\right]} \tag{1.2}$$

I choose the quantile loss function to evaluate the out-of-sample forecasts. Following Giglio, Kelly, and Pruitt (2013), I present the loss using sentiments as a predictor variable relative to the loss from the historical quantile. Equation 1.2 gives this formula, where N is the number of forecasts and ρ_τ is the τth quantile loss function. I assess significance using the Diebold and Mariano (2002) test of superior predictive accuracy of the model with sentiments as a predictor variable given the two sequences of forecast losses.

Table 3 shows the results. Using the ICS as a predictor variable generates a 10%

[4]I select the number of lag lengths with the BIC selection criterion.

forecasting gain relative to using the historical quantile for the 20th percentile of consumption growth innovations. This gain decreases moving to the median and entirely disappears at the 80th percentile[5]. These results suggest that much of the fluctuations in consumer sentiments comes from expectations about downside movements in consumption as opposed to upside movements[6]. The current theoretical literature on news shocks cannot address this asymmetry, but the downturn risk shock can.

4. Innovations in consumer sentiments have an impact on the economy.

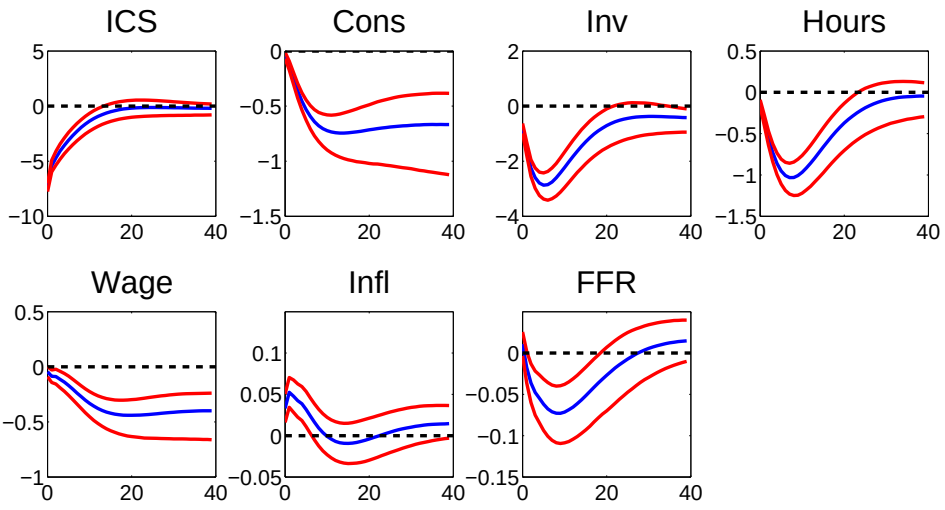


Figure 3 Impulse response to 1 standard deviation decrease in Index of Consumers Sentiments from a Bayesian VAR(4) on quarterly log University of Michigan Index of Consumers Sentiments (ICS), log per capita consumption, log per capita investment, log per capita hours, log per capita real wages, inflation, and federal funds rate. Responses are multiplied by 100. ICS is ordered first and a Cholesky identification strategy is used. Comovements of the impulse responses are robust to ICS ordered last. The blue lines are the pointwise median impulse responses whereas the red lines define the 68% credible set.

Figure 3 shows the responses to a 1 standard deviation negative Cholesky identified

[5]Results are robust to considering the 10th and 90th percentile. I have also tried using alternative definitions of real consumption, deflating the series by the Consumer Price Index less food and energy as well as the Personal Consumption Expenditures price deflator. Full results are in the appendix.

[6]This conclusion is consistent with results by Giglio, Kelly, and Pruitt (2013) that illustrate the predictability of downside innovations in real activity from financial distress measures, as well as the results by Del Negro and Schorfheide (2012) and Alessandri and Mumtaz (2014a) that document the importance of using financial measures to forecast macroeconomic variables across the Great Recession.

shock to innovations in consumer sentiments from a VAR(4) composed of quarterly log University of Michigan Index of Consumers Sentiments, log consumption, log investment, log hours, log real wages, inflation, and federal funds rate. Consumption, investment, hours, and real wages are in per capita values. Consumer sentiments are ordered first. This VAR is similar in spirit to one found in Barsky and Sims (2012). Details of the implementation of the VAR can be found in the appendix.

In response to the decline in sentiments, consumption, investment, and real wages decrease permanently. On impact, consumption and investment both decline. Investment and labor supply both have hump-shaped components. Notice the long-lived response of hours. Almost 25 quarters after an expectations shock, hours still have not fully returned back to steady state. Inflation increases whereas the federal funds rate declines. These responses are all significant, providing suggestive evidence that changes in expectations of agents have important effects on the economy. Moreover, the responses of the variables to the movements in expectations give guidance as to how the model economy should react to expectations shocks.

1.3. Simple model

Given the stylized facts in the previous section, I now turn to illustrating how the downturn risk shock can capture what is found in the data. I start with a real business cycle model with capital and labor to clearly illustrate the properties of the downturn risk shock.

$$\begin{aligned}
&\max_{c_t,i_t,k_{t+1},l_t} \mathbb{E}\sum_{t=0}^{\infty}\beta^t\left(\log\left(c_t-hc_{t-1}\right)-\psi\frac{l_t^{1+\gamma}}{1+\gamma}\right)\\
&s.t.\ c_t+i_t=A_tk_t^{\alpha}l_t^{1-\alpha}\\
&k_{t+1}=i_t+(1-\delta_t)k_t\\
&g_{A,t}=\log A_t-\log A_{t-1}
\end{aligned}\tag{1.3}$$

β	0.994	γ	2
h	0.75	α	0.33
ψ	75.66	δ	0.025

Table 4 Calibration simple economy

The social planner maximizes over consumption c_t, investment i_t, capital k_{t+1}, and labor l_t, subject to the resource constraint and capital accumulation equation. Agents have seperable utility over consumption and labor. They value consumption relative to an internal habit level governed by h. The economy has a Cobb-Douglas production function hit by productivity shocks A_t with growth rate $g_{A,t}$. Depreciation of the capital stock δ_t is stochastic depending upon the state of the world. Stochastic depreciation could come from economic depreciation as in Greenwood, Hercowitz, and Krusell (1997) or economic obsolescence as in Gertler and Kiyotaki (2010), Liu, Waggoner, and Zha (2011), and Gourio (2012). Table 4 gives the calibrated parameters for the simple economy, which are relatively uncontroversial[7]. I consider log-linear approximations to this model economy.

1.3.1. Markov-switching dynamics and the downturn risk shock

There are two observed states of the world, which I characterize as a *downturn state* ($s_t = 1$) and a *normal state* ($s_t = 2$). The downturn state features low productivity growth and high capital depreciation. This state occurs rarely in practice, but agents are aware of its existence and the time-varying risks of transitioning to this state. The normal state has

[7]The parameter δ, to be introduced in the next section, is the steady state depreciation rate.

high productivity growth and low capital depreciation.

$$
\begin{aligned}
& g_{A,t} = (1-\rho_A)\Lambda_A + \rho_A g_{A,t-1} + c_A(s_t) + \epsilon_{A,t}, \epsilon_{A,t} \sim N(0, \sigma_A^2) \\
& \log \delta_t = \log \delta + d(s_t) \\
& c_A(s_t) = \begin{cases} c_A(1) < 0, d(1) > 0 & w.p.\ 1 - p(\overline{s}_{t-1}) \\ c_A(2) > 0, d(2) < 0 & w.p.\ p(\overline{s}_{t-1}) \end{cases} \\
& \mathbb{E}c_A(s), \mathbb{E}d(s) = 0
\end{aligned} \tag{1.4}
$$

Equation 1.4 lays out the productivity growth and depreciation processes. At time t, agents know all variables dated t or earlier. The $g_{A,t}$ process follows an AR(1), as is standard in the literature. The variable $c_A(s_t)$ gives the low growth and high growth values of productivity growth. Likewise, $d(s_t)$ determines the level of capital depreciation across the states. The unconditional means of these parameters are both 0, meaning that they represent deviations around the long-run values. Agents at time t know the probabilities of entering the normal and downturn states in time $t+1$, given by $p(\overline{s}_t)$ and $1 - p(\overline{s}_t)$, respectively. Changes in $p(\overline{s}_t)$, therefore, change the downturn risk that agents face. I introduce $\overline{s}_t$ to index the probabilities with which the economy enters into the normal versus downturn states in the following period. Crucially, notice that a change in this risk does not necessarily have to coincide with a change in observed productivity growth or capital depreciation. It impacts agents' assessments of future prospects, however, and will therefore alter current behavior.

$$
prob\,(\overline{s}_t | \overline{s}_{t-1}) = Q(\overline{s}_{t-1}) \tag{1.5}
$$

Equation 1.5 specifies the dynamics for $\overline{s}_t$. Specifically, the $\overline{s}_t$ states transition according to a first-order Markov process. I allow for many probability states $\overline{s}_t$. This leads to an issue of overparametrization of $p(\overline{s}_t)$ and the transition matrix Q. Having N probability

states means determining $N-1+N(N-1)$ parameters[8]. To solve this issue, I specify the probability process as coming from an N-state Tauchen approximation of the AR(1)-type process specified in equation 1.6.

$$p_t = \frac{e^{\tilde{p}_t}}{1+e^{\tilde{p}_t}} \tag{1.6}$$
$$\tilde{p}_t = (1-\rho_p)\mu_p + \rho_p \tilde{p}_{t-1} + \epsilon_{p,t}, \epsilon_{p,t} \sim N(0, \sigma_p^2)$$

First, I discretize the process $\tilde{p}_t$ into N states. Then, I apply the logistic transformation to each discretized value to constrain each of them onto the unit interval. The Tauchen approximation allows me to jointly determine the probability states and transition matrix as a function of only three parameters: μ_p, ρ_p, σ_p, or the long-run mean, persistence, and standard deviation of the underlying continuous process[9].

Interpretation of downturn transition

The downturn state transition causes a decline in productivity growth and an increase in depreciation. Productivity growth could decline if misallocation increases during downturns. The depreciation increase could capture heightened economic depreciation or obsolescence. Historically, the two have often coincided. Economic obsolescence has been used by Gertler and Kiyotaki (2010) as a driver of financial intermediation disruptions. If the reduced ability of financial intermediation also impacts the economy's ability to allocate capital efficiently, productivity growth could also decline. Economic obsolescence has often been associated with disruptions in financial intermediation. For instance, many consider the recent poor financial conditions driving the downturn in the Great Recession as caused by a large economic depreciation of the housing stock, which tightened bank balance sheets and disrupted lending[10]. Oil supply shocks could also cause productivity growth and economic obsoles-

[8] The $N-1$ comes from the probability values of the N states, while $N(N-1)$ from the transition matrix Q.

[9] Further details of the solution algorithm are in the Structural estimation section of the paper.

[10] Other examples of economic obsolescence leading to financial recessions are the panics of the late 1800s, which were driven by overinvestment and subsequent obsolescence in railroads leading to massive amounts of bank failures.

cence to move together. Changes in oil prices have been known to lead to large reallocations across sectors[11], which could lead to productivity declines if misallocation increases. If the energy shocks lead to abandonment of energy-intensive capital, then economic depreciation would also increase. As this paper focuses on agents' reactions to changing risks of these downturns, I abstract from a further explicit microfoundation of the relationship.

Analysis and simulation of the process

μ_p	2
ρ_p	0.88
σ_p	0.61
$c_A(1)$ $(c_A(2))$	-0.009 (0.002)
$d(1)$ $(d(2))$	1.63 (-0.36)
ρ_A	0.2
σ_A	0.005

Table 5 Downturn risk shocks parameter calibration

Table 5 gives the illustrative calibration for the downturn risk shocks process. A key parameter of interest is μ_p, the long-run mean parameter. The sign of μ_p determines whether the process captures downturn risk or upturn risk. The logistic transformation ensures that a value $\tilde{p}_t = 0$ gives exactly one half chance of transitioning to either state at time $t+1$. Values of $\tilde{p}_t > 0$ mean a higher probability of transitioning to the normal state while $\tilde{p}_t < 0$ imply the opposite. A long-run mean of $\mu_p > 0$ implies that the $\tilde{p}_t$ process should spend more time above 0, leading to a higher probability on average of transitioning to the normal state. Unconditional expectations of $c_A(s_t)$ and $d(s_t)$ both equalling 0 then imply that $|c_A(1)| > |c_A(2)|, |d(1)| > |d(2)|$. This parametrization captures the intuition of a larger magnitude, but relatively infrequent, downturn state and a smaller magnitude, but more frequent, normal state. Moreover, a higher value of μ_p implies a larger asymmetry. For

[11] See for example Davis and Haltiwanger (2001).

the illustrative calibration, I set the value $\mu_p = 2$, which leads to $c_A(s_t)$ and $d(s_t)$ having a rather large degree of asymmetry in deviations from steady state[12].

The asymmetries in the process generate unconditional negative skewness for the productivity growth and positive skewness for the depreciation rate processes. For the calibration above, the skewness of productivity growth is -0.25 while that of depreciation is 1.63. This asymmetry translates to asymmetry in the endogenous variables as well, to be discussed in a following subsection. Therefore, the model can capture the asymmetry of the business cycle. Deep recessions are associated with more frequent transitions to the downturn state - or downside shocks - causing rapid decreases in the endogenous variables. Because the probability of the downturn state occurring, given by $p(\overline{s}_t)$, is known by agents, the downturn risk shock can capture the stylized fact that changes in agents' expectations are associated with downside innovations in the economy.

The parameters ρ_p and ρ_A govern the persistence of the probability and productivity growth processes, respectively. A higher value of ρ_p makes the probability process more persistent, meaning that an increased probability of transitioning to the downturn state at time t leads to higher probabilities of a downturn state for an extended period. Increased ρ_A leads to a longer propagation of realized productivity growth changes.

The interaction of the two parameters, however, can lead to many interesting shapes for expected productivity growth as a result of a change in downturn risk. Equation 1.7, the impulse response to a probability shock at time t, illustrates why this happens (Call $\hat{g}_{A,t+k}$ the deviation of $g_{A,t+k}$ from its long-run mean.). An innovation to the probability process does not change productivity growth at period t. Increasing the probability of transitioning to the downturn state, for example, means that at time $t+1$, there exists a higher probability of a downturn state transition, lowering $E_t g_{A,t+1}$. One can think of this as time $t+1$ news.

[12]The exact unconditional probabilities of being in a downturn state versus a normal state depend on all 3 parameters of the probability process as well as the number of points of discretization. For a 10-state discretization, on average, the process spends 82 percent of the time in the normal state and 18 percent of the time in the downturn state. As the number of discretization points increase, of course, the impact of the number of discretization points diminishes.

Moving forward, there now exists two effects on $E_t g_{A,t+k}$. First, if $\rho_p > 0$, there still exists a higher probability of a transition to the downturn state, as the probability process has persistence. Think of this as time $t+k$ news. Second, with $\rho_A > 0$, the previous period's conditional expectation of the process propagates forward because the $g_{A,t}$ process itself is persistent as well. The first effect feeds into the second one, implying that a probability shock can persist for quite some time. In fact, if ρ_A and ρ_p are persistent enough, the impulse response of $g_{A,t}$ can even feature a hump-shape, increasing in the initial periods before mean-reverting, a feature that innovations in a standard real business cycle model cannot achieve.

$$
\begin{aligned}
&\hat{g}_{A,t} = 0 \\
&\mathbb{E}_t \hat{g}_{A,t+1} = \underbrace{(1 - p_t^{(1)}) c_A(1) + p_t^{(1)} c_A(2)}_{t+1 \text{ news}} \\
&\mathbb{E}_t \hat{g}_{A,t+2} = \rho_A((1 - p_t^{(1)}) c_A(1) + p_t^{(1)} c_A(2)) + \underbrace{(1 - p_t^{(2)}) c_A(1) + p_t^{(2)} c_A(2)}_{t+2 \text{ news}} \ ... \\
&\mathbb{E}_t \hat{g}_{t+k} = \sum_{j=1}^{k} \rho_A^{k-j} \left((1 - p_t^{(j)}) c_A(1) + p_t^{(j)} c_A(2) \right) \\
&p_t^{(n)} = \mathbb{E}_t p(\overline{s}_{t+n-1})
\end{aligned}
\tag{1.7}
$$

Figure 4 shows impulse responses of productivity growth to a decrease in $p(\overline{s}_t)$, which makes the downturn state more likely. As both ρ_p and ρ_A are positive, expected productivity growth responds in a hump-shaped fashion, increasing in the first few periods before declining back to its steady state value. For comparison purposes, the red line plots an unanticipated productivity growth innovation of magnitude around the same as the maximal productivity growth response to the probability shock. Relative to the productivity growth response from the probability innovation, following an unanticipated innovation, productivity growth quickly declines back to its steady state value. The longer productivity growth response from the probability shock comes from the second source of persistence, namely the positive autocorrelation of the probability process. In this way, the down-

turn risk shock naturally captures the longer-horizon, medium-frequency, fluctuations that Comin and Gertler (2006) described as a property of the medium-frequency cycle.

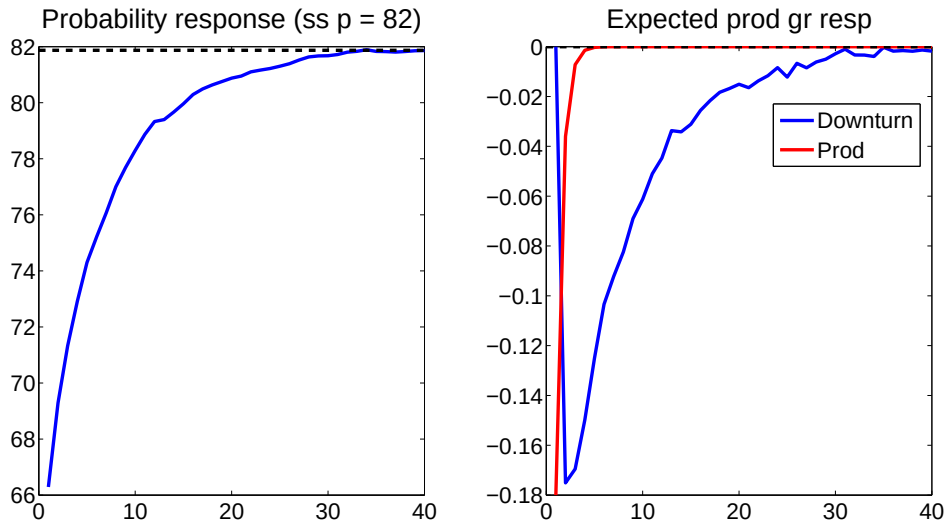


Figure 4 Impulse response of productivity growth to downturn risk shock (blue) compared to unanticipated productivity growth innovation (red) (multiplied by 100)

Along with the movements from productivity growth, the downturn risk shock also results in changes to the expected depreciation rate, as shown in figure 5. Expected depreciation reaches its highest value the period after the probability innovation, representing the time period of maximal risk of transitioning to the high depreciation state, and reverts back to steady state in an AR(1) fashion. Allowing for some persistence in the realizations of the depreciation shock produces a response analagous to that of productivity growth. Therefore, the downturn risk shock also provides news about depreciation rate movements.

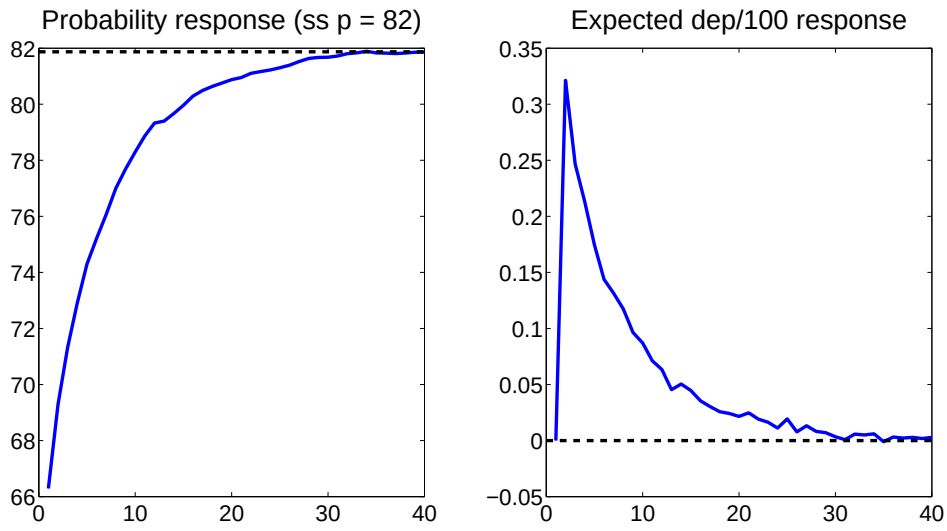


Figure 5 Impulse response of depreciation rate to downturn risk shock (multiplied by 100 log deviations from steady state)

1.3.2. Endogenous responses of agents to a downturn risk shock and a comparison to news shocks

A well-known stylized fact of business cycle fluctuations is the comovement between output, consumption, and investment over the cycle. To empirically qualify as a main driving force of fluctuations, a shock must produce such comovement. Moreover, the VAR evidence suggests that a negative shock to consumer sentiments generates declines in consumption and investment. It is well known since Beaudry and Portier (2004) that expectations shocks, the category in which the downturn risk shock falls, have difficulty generating such comovement. In this section, I discuss the model in equation 1.3, which can lead to comovement in response to the downturn risk shock. At the same parameter values, in response to a variety of news shocks, the model does not generate comovement.

It is instructive to compare the agents' responses to the downturn risk shock with those to the news shock. I follow the specification in Schmitt-Grohe and Uribe (2012) in allowing for $m-$period ahead news in productivity growth, MEI [13], and depreciation, shown in

[13]The MEI shock (μ_t) alters the efficiency with which investment transforms into capital: $k_{t+1} = i_t\mu_t +$

equation 1.8. The literature has been largely concerned with generating comovement from productivity news. Jaimovich and Rebelo (2009) present minimal modifications to the real business cycle model to generate procyclical movements in response to productivity and investment-specific technology news. Analyzing the downturn risk shock's relationship to news in both productivity growth and the capital accumulation equation proves useful, as the downturn risk shock alters expectations in both areas of the economy. Since the downturn risk shock has a depreciation news component, I also compare the response of the economy to depreciation news. In comparing the downturn risk shock to the three news shocks, I emphasize that it is not simply the presence of news in both sectors of the economy that drive the results, but rather the interaction between the two. Moreover, none of the news shocks can capture the asymmetric risks of the downturn risk shock.

I now introduce a concept called the *pure downturn risk shock*. It is important to distinguish this exercise from the downturn risk shock. A downturn risk shock changes the probability with which a transition to the downturn state occurs. There are two effects from the downturn risk shock that impact allocations moving forward. First, agents' expectations change. Second, the frequency of transitions to the downturn state changes. Both change allocations, but it is interesting to isolate the impact on allocations of the change in agents' expectations. Consider the following thought experiment: suppose the risk of the downturn state increased, but the realized sequence of productivity growth and depreciation values stayed constant relative to a scenario where downturn risk did not change. Agents, in response to the change in downturn risk, would still change their allocations based on their differing expectations of the future. I call this scenario a *pure downturn risk shock*. This isolates the expectations effect coming from the change in downturn risk abstracting away

$(1 - \delta_t)k_t$.

from any movements in observed fundamentals.

$$g_{A,t} = \rho_A g_{A,t-1} + \epsilon_{A,t} + \epsilon^m_{A,t-m}, \epsilon_{A,t} \sim \mathcal{N}(0, \sigma^2_A), \epsilon^m_{A,t-m} \sim \mathcal{N}(0, \sigma^2_{A,m}) \tag{1.8}$$

$$\log \mu_t = \rho_\mu \log \mu_{t-1} + \epsilon_{\mu,t} + \epsilon^m_{\mu,t-m}, \epsilon_{\mu,t} \sim \mathcal{N}(0, \sigma^2_\mu), \epsilon^m_{\mu,t-m} \sim \mathcal{N}(0, \sigma^2_{\mu,m}) \tag{1.9}$$

$$\log \delta_t = \log \delta + \epsilon_{\delta,t} + \epsilon^m_{d,t-m}, \epsilon_{d,t} \sim \mathcal{N}(0, \sigma^2_d), \epsilon^m_{d,t-m} \sim \mathcal{N}(0, \sigma^2_{d,m}) \tag{1.10}$$

ρ_A	0.2	ρ_μ	0.5
$\sigma_{A,m}$	0.005	$\sigma_{A,m}$	0.05
$\sigma_{d,m}$	0.05		

Table 6 News shocks parameter calibration ($m = 4$)

The pure downturn risk shock has a connection with an unrealized news shock. In the latter case, agents receive news about the future that does not realize. The exercise involves having an unanticipated shock $\epsilon_{A,t}, \epsilon_{\mu,t}, \epsilon_{d,t}$ cancel out the arrival of news. Despite the similarities, a crucial difference is that the pure downturn risk shock is not a measure zero event. Because a pure downturn risk shock only involves a change in probabilities of future states, it makes sense to think about the scenario of a change in risk but no difference in realized productivity growth and depreciation. This allows for one to meaningfully econometrically identify historical situations best described by a change in the risk profile faced by the agents but no change in observed fundamentals. On the other hand, an unrealized news shock is a measure zero event, and therefore can only be thought of as a theoretical construct.

For the analytical analysis, however, I consider pure downturn risk shocks and unrealized news shocks. Doing so isolates the expectations effects on allocation decisions.

Consider the consumption and investment responses of the agents to a pure downturn risk shock, shown in figure 6. An increase in the downturn risk probability causes agents to decrease consumption and investment, generating comovement. Because both consumption

and investment decrease, output must decrease as well. This comovement occurs with the introduction of a moderate amount of habit persistence $h = 0.75$. The lower productivity growth expectations lead agents to lower consumption immediately due to the wealth effect. Because productivity does not change, this usually puts strong upward pressure on investment from the resource constraint. In this case, investment declines due to the expected increase in the depreciation rate, which lowers the expected return on investment. Consumption habit persistence also helps reduce the immediate incentive to cut consumption, as agents do not want to change consumption allocations too quickly, thereby aiding in generating comovement.

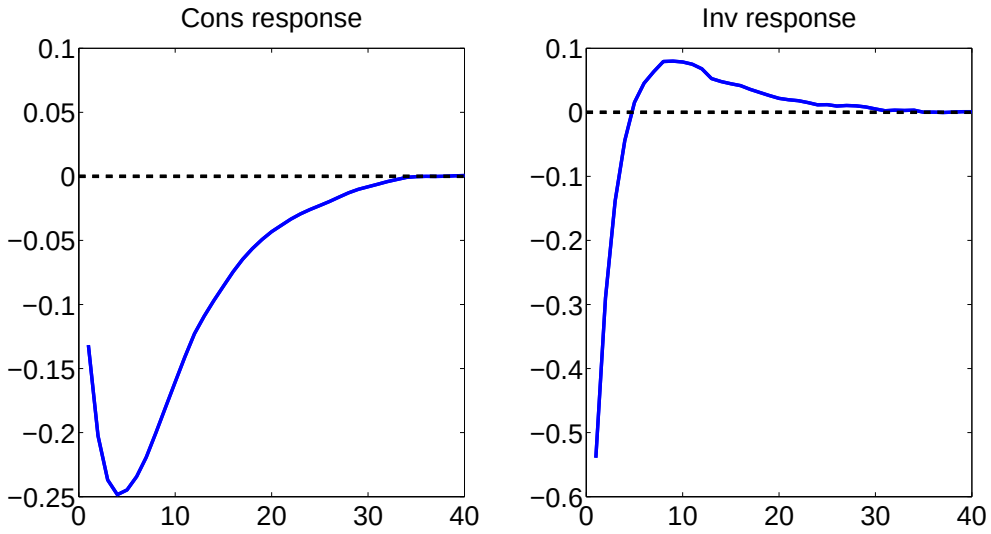


Figure 6 Impulse responses of consumption and investment to pure downturn risk shock (multipled by 100 log deviations from steady state)

Figure 7 displays the response of agents to the three 4−period ahead news shocks at the same parameter calibration where appropriate. The blue line gives the response to negative productivity growth news. Having only consumption habits cannot produce comovement in response to productivity growth news. Agents, understanding that productivity will decrease in the future, would like to immediately cut consumption. This leads them to increase investment. Consumption habits mitigate the incentive to immediately decrease consumption, but as can be seen from the impulse response, it is not enough to generate

comovement. Relative to productivity growth news, the downturn risk shock also increases the expected depreciation rate, thereby reducing the incentive to invest.

The endogenous response to negative MEI news (black line) has the same difficulty in generating comovement. In this case, agents understand that the future efficiency of investment will decrease. When the news shock realizes, agents would want to cut investment, thereby increasing consumption from the resource constraint. Due to habit persistence, they would not want to change consumption allocations too quickly. Therefore, in response to expectations of a future decline in MEI, agents decide to immediately begin increasing consumption. Increased consumption tends to produce decreased investment from the resource constraint, and the news shock does not lead to comovement.

News about future increases in depreciation (red line) generate an immediate decline in investment and corresponding increase in consumption. Investment declines because the knowledge of future depreciation rate increases makes investment today less attractive. Again, due to the resource constraint holding, decreased investment tends to increase consumption. The downturn risk shock counteracts this issue by having a productivity growth component. Decreased productivity growth expectations incentivize agents to decrease consumption and increase investment.

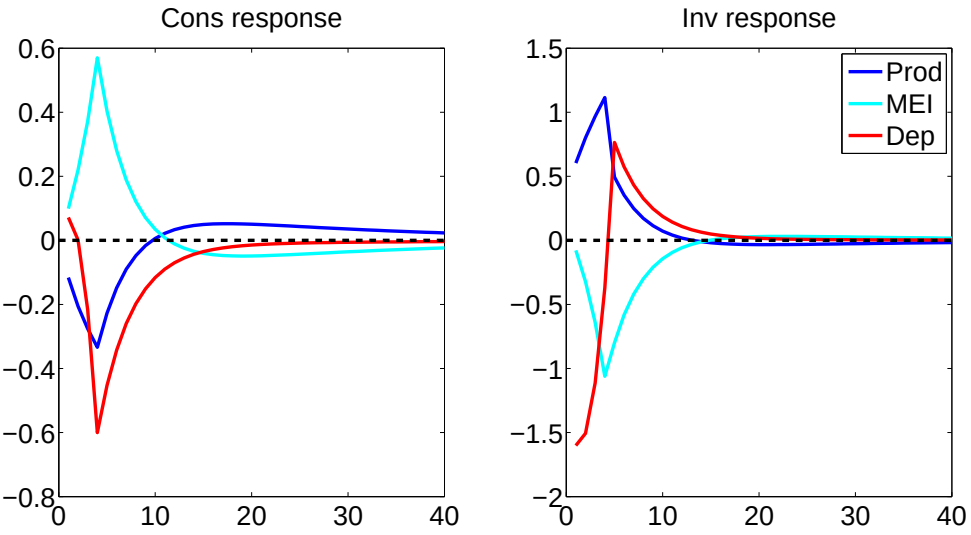


Figure 7 Impulse responses of consumption and investment to various negative 1 standard deviation 4−period ahead unrealized news shocks (multipled by 100 log deviations from steady state)

One more important difference in comparing the impulse responses in figures 6 and 7 involves their dynamics beyond the immediate response. The downturn risk shock does a relatively good job at producing reasonable consumption and investment dynamics. They tend to both decline and increase together, which matches the general feature of the business cycle. This happens because the downturn risk shock simultaneously contains news in both sectors, so the agents not only want to decrease consumption from productivity growth news, but they also want to cut investment from depreciation news. The news shocks, on the other hand, lead to a correlation between consumption and investment of close to -1. In other words, when consumption increases, investment tends to decrease and vice versa.

With further modifications to the model involving stronger adjustment costs, it is of course possible to produce comovement in response to all three news shocks. The interplay between productivity growth and depreciation that the downturn risk shock features, however, reduces the need for such adjustment costs.

1.3.3. Why both productivity growth and depreciation are necessary

Are both productivity growth and depreciation necessary for the downturn risk shock to produce comovement between consumption and investment without large adjustment costs? I answer this question by looking at the responses of the economy first shutting down productivity growth and then removing depreciation, presented in figure 8.

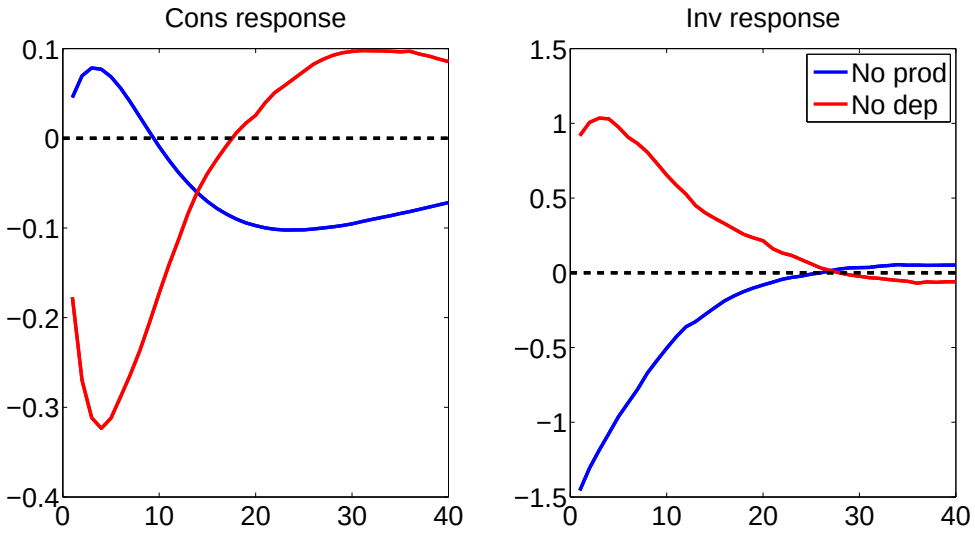


Figure 8 Impulse responses of consumption and investment to pure downturn risk shock shutting down productivity growth and depreciation (multipled by 100 log deviations from steady state)

With productivity growth shut down, the pure downturn risk shock causes an increase in consumption and decrease in investment. Agents expect higher depreciation in the future, which lowers investment. From the resource constraint holding, this leads to an upward response in consumption. Shutting down depreciation leads to the opposite situation. Now agents decrease consumption due to expectations of future productivity growth declines, which leads to an increase in investment. The interplay between the two effects creates the comovement.

1.3.4. Transition to downturn state

It is important to also analyze the response of the economy to a realization of the downturn state. The asymmetry of the downturn state generates the negative skewness seen in the data. As large downward movements in output, consumption, and investment tend to occur together in the data, it is important to have the proper comovement following a transition to the downturn state.

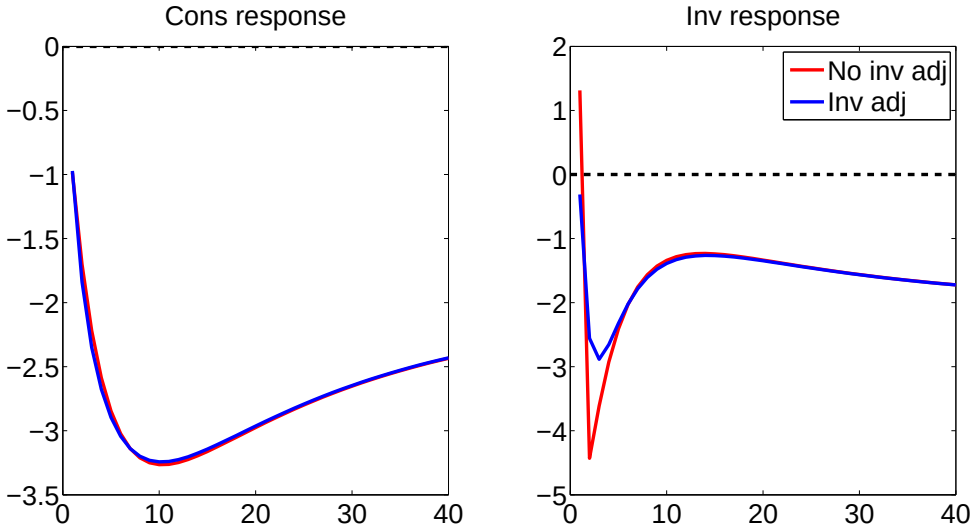


Figure 9 Impulse responses of consumption and investment to realization of downturn state with (blue) and without (red) investment adjustment costs ($S'' = 1$) (multipled by 100 log deviations from steady state)

Figure 9 shows the response of the economy with and without a small amount of quadratic investment adjustment costs[14]. Investment adjustment costs are just one of many ways to generate comovement between consumption and investment in response to a realization of the downturn state. Consumption drops immediately due to the permanent nature of the productivity decrease. Having a depreciation component incentivizes agents to increase investment during the period of the transition because they know the capital stock will

[14]Investment adjustment costs alter the capital accumulation equation as follows: $k_{t+1} = (1-\delta_t)k_t + \left(1 - \frac{S''}{2}\left(\frac{i_t}{i_{t-1}} - e^{\Lambda_z}\right)^2\right) i_t$.

decrease by a large amount in the next period. As the capital stock decreases in the following period, thereby decreasing output in the economy, investment then drops by a large amount. Including investment adjustment costs, which penalizes for large movements in investment, encourages agents to begin decreasing investment the period of the downturn state transition.

Simulating the economy[15] in equation 1.3 at the calibrated parameters leads to skewness in output growth of -0.56, consumption growth of -0.96, and investment growth of -0.23. In contrast, the model with only normally distributed shocks would not produce any skewness in these variables, and thus could not capture the rapid decreases the model augmented with downturn shocks can generate.

1.4. Structural estimation

The previous section illustrated the downturn risk shock's ability to simultaneously capture the stylized facts in section 2. An important remaining question is the empirical importance of the downturn risk shock in a larger scale model with competing structural shocks. Are the properties of the downturn risk shock empirically relevant? I use Bayesian methods to estimate a dynamic equilibrium model in answering these questions. I first present the full quantitative model. Then, I lay out the nonstandard model solution method. Third, I present my estimation algorithm and data. Finally, I discuss empirical results.

1.4.1. Model

I consider a New Keynesian model to conduct my analysis. The model follows the tradition of Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007). It includes consumption habits, price, nominal wage, investment, and capacity utilization adjustment costs, and a Taylor rule that responds to the inflation rate and output growth. Eight shocks drive the model: productivity growth, productivity growth news, marginal efficiency

[15] Simulation of length 25000.

of investment, labor supply, inflation target, government spending, monetary policy, and downturn risk. The model also has two states of the world: the downturn and normal states.

Household

There exists a representative household that has the following preferences:

$$\mathbb{E}\sum_{t=0}^{\infty}\beta^t\left(\log(c_t - hc_{t-1}) - \psi_t\int_0^1 \frac{l_{j,t}^{1+\gamma}}{1+\gamma}dj\right) \tag{1.11}$$

The parameter β is the discount factor. The household has separable utility over consumption and labor. Consumption is valued relative to an internal habit level with the parameter h determining the amount of habit persistence. The household supplies a unit mass of differentiated labor to the market. A labor supply shock ψ_t perturbs the household's utility cost of supplying labor.

The household maximizes with respect to consumption (c_t), real wages ($w_{j,t}$), one-period bonds (b_{t+1}), investment (i_t), capital (k_{t+1}), and capacity utilization (u_t). Its budget constraint consists of:

$$\begin{aligned} &c_t + i_t + \int_0^1 \frac{\phi_w}{2}\left(\frac{w_{j,t}}{w_{j,t-1}}\frac{\Pi_{t+\tau}}{\Pi^*_{t+\tau}} - e^{\Lambda_z}\right)^2 y_t d_j + \frac{b_{t+1}}{p_t} = \\ &\int_0^1 w_{j,t}l_{j,t}dj + r_t u_t k_t + R_{t-1}\frac{b_t}{p_t} + T_t + F_t \end{aligned} \tag{1.12}$$

The household sets real wage rates $w_{j,t}$ subject to Rotemberg (1982) quadratic adjustment costs. The nominal wage adjustment costs are relative to steady state real wage growth e^{Λ_z} and inflation target Π^*_t. By investing in bonds b_{t+1}, the household earns a nominal return R_t. Renting out capital earns a return of r_t. Finally, government transfers T_t and firm profits F_t are given to the households lump-sum. The price level in the economy is given by p_t.

Labor market

Labor used by intermediate goods producers gets aggregated by a labor packer with the production function:

$$l_t = \left(\int_0^1 l_{j,t}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}} \tag{1.13}$$

The intermediate goods producers pay the real wage rate w_t for one unit of aggregate labor l_t.

Households face the following maximization problem, where $\lambda_{t+\tau}$ is the Lagrange multiplier on the budget constraint at time $t+\tau$:

$$\max_{w_{j,t}} \mathbb{E}_t \sum_{\tau=0}^{\infty} \beta^\tau \left\{ -\psi_t \frac{l_{j,t+\tau}^{1+\gamma}}{1+\gamma} + \lambda_{t+\tau} \left(w_{j,t+\tau} l_{j,t+\tau} - \frac{\phi_w}{2} \left(\frac{w_{j,t+\tau}}{w_{j,t+\tau-1}} \frac{\Pi_{t+\tau}}{\Pi^*_{t+\tau}} - e^{\Lambda_z} \right)^2 y_t \right) \right\} \tag{1.14}$$

$$s.t. \quad l_{j,t+\tau} = \left(\frac{w_{j,t+\tau}}{w_{t+\tau}} \right)^{-\eta} l_{t+\tau}$$

They optimally set wages, taking into account the adjustment costs, to balance out the negative utility flow of working with the positive flow of increased resources.

Capital accumulation

Capital accumulates according to the law of motion in equation 1.15 and is subject to quadratic adjustment costs. Moreover, increased utilization leads to higher depreciation of the capital stock as in Greenwood, Hercowitz, and Huffman (1988), Jaimovich and Rebelo (2009), and Schmitt-Grohe and Uribe (2012).

$$k_{t+1} = (1 - \delta(d_t, u_t))k_t + \mu_t \left(1 - \frac{S''}{2} \left(\frac{i_t}{i_{t-1}} - e^{\Lambda_z} \right)^2 \right) i_t \tag{1.15}$$

$$\delta(d_t, u_t) = \delta e^{d(s_t)} + \Phi_1(u_t - 1) + \frac{\Phi_2}{2}(u_t - 1)^2 \tag{1.16}$$

Depending on the state of the world s_t, depreciation can either take on a high value $d(1) > 0$ or low value $d(2) < 0$. The model has both shocks to depreciation and the marginal efficiency of investment μ_t [16]. I have already discussed the interpretation of the downturn state when introducing the downturn risk shock. The marginal efficiency of investment shock captures changes in the efficiency with which the economy converts investment goods into the capital stock. Justiniano, Primiceri, and Tambalotti (2011a) also take this interpretation.

Firms

A final goods producer aggregates intermediate goods production into the final good y_t and sells it at price p_t:

$$y_t = \left(\int_0^1 y_{i,t}^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \tag{1.17}$$

Intermediate goods producers have a Cobb-Douglas production function subject to a productivity shock A_t:

$$y_{i,t} = A_t k_{i,t}^{\alpha} (l_{i,t})^{1-\alpha} \tag{1.18}$$

When setting prices, intermediate goods producers face quadratic adjustment costs, which leads to the maximization problem:

$$\max_{p_{i,t}} \mathbb{E}_t \sum_{\tau=0}^{\infty} \beta^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} \left(\left(\frac{p_{i,t+\tau}}{p_{t+\tau}} - mc_{t+\tau} \right) \left(\frac{p_{i,t+\tau}}{p_{t+\tau}} \right)^{-\epsilon} y_{t+\tau} - \frac{\phi_p}{2} \left(\frac{p_{i,t+\tau}}{\Pi^*_{t+\tau} p_{i,t+\tau-1}} - 1 \right)^2 y_{t+\tau} \right) \tag{1.19}$$

Intermediate goods producers set prices to maximize future profits discounted according to the pricing kernel of the household. If the firm sets price $p_{i,t}$ such that it grows according to the time-varying inflation target of the central bank (Π^*_t), then it does not face any adjustment cost.

[16] Saijo (2013) also considers an environment with both depreciation and marginal efficiency of investment shocks.

Government

Government spending g_t evolves as a stochastic fraction of output y_t according to the process $\tilde{g}_t$. Fiscal policy is Ricardian and financed by one-period bonds.

$$g_t = \left(1 - \frac{1}{\tilde{g}_t}\right) y_t \tag{1.20}$$

The monetary authority sets monetary policy according to a Taylor rule with a time-varying inflation target Π_t^*:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi_t^*}\right)^{\gamma_\Pi} \left(\frac{\frac{y_t}{y_{t-1}}}{\Lambda_y}\right)^{\gamma_y}\right)^{1-\gamma_R} e^{\epsilon_{r,t}} \tag{1.21}$$

The parameter γ_R determines the amount of interest rate smoothing present in the policy rule. Parameters γ_Π and γ_y determine the aggressiveness of central bank responses to inflation deviations from target inflation and output growth deviations from steady state growth. I impose $\gamma_\Pi > 1$ and $\gamma_y > 0$ to ensure determinacy. The random variable $\epsilon_{r,t}$ is a monetary policy shock.

Aggregation and market clearing

Equation 1.22 gives the aggregate resource constraint, with consumption, investment, government spending, price, and wage adjustment costs equaling total output. Factor market clearing ensures that total labor employed by firms equals the amount supplied by households and total capital rented equals the utilization-adjusted capital stock.

$$c_t + i_t + g_t + \frac{\Phi_w}{2}\left(\frac{w_t}{w_{t-1}}\frac{\Pi_t}{\Pi_t^*} - e^{\Lambda_z}\right)^2 y_t + \frac{\Phi_p}{2}\left(\frac{\Pi_t}{\Pi_t^*} - 1\right)^2 y_t = y_t \tag{1.22}$$

$$\int_0^1 l_{j,t} dj = l_t \tag{1.23}$$

$$u_t k_t = \int_0^1 k_{i,t} di \tag{1.24}$$

Exogenous processes

In total, the baseline model has eight exogenous shocks and two states of the world. The productivity shock is nonstationary.

$$g_{A,t} = \log A_t - \log A_{t-1} \tag{1.25}$$

I allow for both the downturn risk shock and 4−period ahead news in the productivity growth process, along with the standard unanticipated innovation. The news shock is the primary competitor to the downturn risk shock in that it also impacts agents' expectations without changing contemporaneous observed productivity growth.

The marginal efficiency of investment shock is stationary in logs:

$$\log \mu_t = \rho_\mu \log \mu_{t-1} + \epsilon_{\mu,t}, \epsilon_{\mu,t} \sim \mathcal{N}(0, \sigma_\mu^2) \tag{1.26}$$

The labor supply shock is stationary in logs around steady state level $\log \psi$:

$$\log \psi_t = (1 - \rho_\psi) \log \psi + \rho_\psi \log \psi_{t-1} + \epsilon_{\psi,t}, \epsilon_{\psi,t} \sim \mathcal{N}(0, \sigma_\psi^2) \tag{1.27}$$

The inflation-target shock is stationary in logs around steady state level $\log \Pi$:

$$\log \Pi_t^* = (1 - \rho_{\Pi^*}) \log \Pi + \rho_{\Pi^*} \log \Pi_{t-1}^* + \epsilon_{\Pi^*,t}, \epsilon_{\Pi^*,t} \sim \mathcal{N}(0, \sigma_{\Pi^*}^2) \tag{1.28}$$

The government spending shock is also stationary in logs around steady state level $\log \tilde{g}$.

$$\log \tilde{g}_t = (1 - \rho_g) \log \tilde{g} + \rho_g \log \tilde{g}_{t-1} + \epsilon_{g,t}, \epsilon_{g,t} \sim \mathcal{N}(0, \sigma_g^2) \tag{1.29}$$

Finally, the monetary policy shock is iid:

$$\epsilon_{r,t} \sim \mathcal{N}(0, \sigma_r^2) \tag{1.30}$$

1.4.2. Solution method

I consider a log-linear solution to the model and use the perturbation methodologies on solving dynamic equilibrium models with Markov-switching. Specifically, I show how to transform the model into one of the form in equation 1.31 where ζ_t are the known variables to the agents as of time t, ϵ_t are the structural innovations, η_t are expectational errors, and $\{\Gamma_0, \Gamma_1, \Gamma_c(\tilde{s}_t), \Psi, \Pi\}$ are functions of structural parameters with $\Gamma_c(\tilde{s}_t)$ switching according to the state $\tilde{s}_t$.

$$\Gamma_0 \zeta_t = \Gamma_c(\tilde{s}_t) + \Gamma_1 \zeta_{t-1} + \Psi \epsilon_t + \Pi \eta_t \tag{1.31}$$

Solution algorithm:

1. Perform an N-state Tauchen (1986) approximation to the probability process.

   The probability process p_t in equation 1.6 is a nonlinear transformation of the AR(1) process $\tilde{p}_t$. The Tauchen approximation discretizes the continuous $\tilde{p}_t$ process into a finite-state Markov chain. Applying the nonlinear transformation to each discretized point then gives the finite-state approximation to the process p_t.

2. Expand the 2-state Markov-switching process with time-varying probabilities into a $2N$-state Markov-switching process with constant transition probabilities (transition matrix Q) where each new state is indexed by the mean state and probability state.

The Tauchen approximation produces N probability states. These probability states are independent of the two mean states. Therefore, combining all combinations of mean and probability states gives $2N$ possible states ($\tilde{s}_t$). Call the expanded transition matrix between these states Q. Q is formed by appropriately combining the transition matrix from the Tauchen approximation and the transition probabilities between the mean states. Switches in the mean state change the realized values of the productivity and depreciation processes while switches in the probability states change the mean state probability transitions moving forward. The states switch in an analogous fashion to Schorfheide (2005) (Bianchi and Melosi (2013) present an example with productivity growth mean switches depending upon the state). These enter into the log-linearized equilibrium conditions as switches in the constant term $\Gamma_c(\tilde{s}_t)$[17]. The log-linearized equilibrium can now be represented as in equation 1.31.

3. Because the Markov-switching enters linearly into the log-linearized solution of the model, use solution techniques developed in Schorfheide (2005).

   Solving for decision rules when the equilibrium conditions are in the form of equation 1.31 can be done using solution methods for linear rational expectations models.

Relationship with Foerster, Rubio-Ramirez, Waggoner, and Zha (2013)

Foerster, Rubio-Ramirez, Waggoner, and Zha (2013) present a general framework for using perturbation methods to solve dynamic equilibrium models with Markov-switching. The authors advance the partition perturbation method, which distinguishes between Markov-switching parameters that affect the steady state (in their, notation, called $\theta_1(k)$), and Markov-switching parameters that do not affect the steady state ($\theta_2(k)$). For parameters that do affect the steady state, they define the perturbation function as in equation 1.32, where n_s is the total number of states and $\overline{\theta}_1$ are the steady states of the Markov-switching parameters defined as the means of the ergodic distributions while $\hat{\theta}_1$ are deviations from

[17]The full log-linearized equilibrium conditions can be found in the appendix.

the steady states:

$$\theta_1(k) = \chi\hat{\theta}_1(k) + (1-\chi)\overline{\theta}_1 \tag{1.32}$$
$$1 \leq k \leq n_s$$

On the other hand, for parameters that do not effect the steady state, they do not perturb the parameters. Not perturbing the Markov-switching parameters in $\theta_2(k)$ allows for potentially time-varying Taylor series coefficients in the solution depending on the state. It also creates a difficult problem of solving n_s systems of quadratic equations. Foerster, Rubio-Ramirez, Waggoner, and Zha (2013) provide a methodology using Grobner bases to solve the model in this situation and show that not perturbing parameters in $\theta_2(k)$ generates more accurate solutions.

In the context of my model, $\theta_2(k)$ is empty, as the only parameters within the model that switch - $c_A(k)$, $d(k)$ - affect the steady state. Therefore, the Taylor series coefficients do not switch depending on the state. Moreover, in the case of $\theta_2(s_t)$ empty, the more advanced algorithm collapses to my method.

1.4.3. Estimation

I use quarterly data of output growth, consumption growth, investment growth, hours, inflation, federal funds rate, and real wage growth from $1960Q2 - 2011Q4$ to estimate the model. This set of data contains the standard macroeconomic series commonly used in estimations of New Keynesian dynamic equilibrium models. I set the starting date to $1960Q2$ to match the beginning of my ICS data, as I will use that data as external validation of the model.

I estimate the following parameters in the model:

Exogenous shocks: $\rho_A, \rho_\mu, \rho_{\pi^*}, \rho_g, \rho_\psi, \sigma_A, \sigma_{A,4}, \sigma_\mu, \sigma_{\pi^*}, \sigma_g, \sigma_r, \sigma_\psi$

Regime-switching: $c_A(1), d(1), \rho_p, \sigma_p, \mu_p$

Structural: $S'', \phi_p, \phi_w, \gamma_R, \gamma_\pi, \gamma_y, h$

These include all parameters of the productivity growth, MEI, inflation target, government spending, monetary policy, labor supply, and depreciation shock processes, adjustment cost parameters, Taylor rule parameters, and the habit persistence value[18].

I fix the rest of the parameters at the values defined in table 7. Most parameters are fixed at standard values in the literature. I set N, the number of probability states in the Tauchen approximation, to 10.

α	0.33	β	0.994
Λ_A	0.0026	Π	0.0092
$\tilde{g}$	0.19	δ	0.025
ϵ	21	η	21
ψ	75.66	γ	2
Φ_2	0.01		

Table 7 Fixed parameters

I use an adaptive random-walk Metropolis-Hastings algorithm with 80000 draws (burn-in 50000) to sample from the posterior density. Forming the posterior requires constructing the likelihood and prior. Because the model has Markov-switching states, I use the Kim (1994) filter to approximate the likelihood as is standard in the literature on estimating dynamic equilibrium models with Markov-switching. The Kim filter is an extension of the

[18] I estimate the model using simulated data to check for identification issues and do not find any. The results are in the appendix.

Hamilton (1989) filter of Markov-switching processes to the linear state space model[19].

I adopt relatively standard priors in the literature, largely following the specifications of Guerron-Quintana (2010). The priors on the parameters for the downturn risk shock deserve more discussion. I put flat priors on all these parameters. They include the Markov-switching components in the productivity growth and depreciation processes ($c_A(s_t)$, $d(s_t)$) as well as the parameters in the probability process (ϕ_p, μ_p, σ_p). I therefore let the data speak as far as the estimated values of these parameters. The full list of priors can be found in the appendix.

1.4.4. Empirical results

Estimated probability process

The estimated probability process is highly persistent, as can be seen in table 8. Moreover, relative to the standard deviation of the unanticipated shock, the transition to the downturn state has a large effect on the productivity growth process. The economy spends on average 74% of the time in the normal state and 26% of the time in the downturn state at the posterior mode estimated parameters.

[19] An alternative route would be to use the particle filter for conditionally Gaussian models. A full discussion can be found in Herbst and Schorfheide (2014).

Parameter	Lower (10%)	Mean	Upper (90%)
ρ_A	0.47	0.53	0.59
σ_A	4.5×10^{-3}	5.1×10^{-3}	5.8×10^{-3}
$c_A(1)$	-0.01	-8.7×10^{-3}	-6.1×10^{-3}
$c_A(2)$	1.7×10^{-3}	2.3×10^{-3}	2.9×10^{-3}
$d(1)$	0.71	0.99	1.22
$d(2)$	-0.34	-0.26	-0.20
ρ_p	0.85	0.87	0.89
μ_p	0.84	1.34	1.70
σ_p	0.63	0.75	0.93

Table 8 Probability shock estimated parameters posterior intervals

Figure 10 contains two graphs. The upper graph displays smoothed estimates of the realized state at the posterior mode parameters. If the blue bar is closer to 1, the estimates place a higher weight on the downturn state whereas if the blue bar is closer to 0, the estimates place a higher weight on the normal state. Especially across the decade of the 1970s and early 1980s, with the OPEC oil embargo and poor policy by the Federal Reserve, the economy was hit by many large downside innovations. The Great Moderation had fewer transitions to the downturn state, especially in the latter halves of the 1980s and 1990s decades. From 2004 onwards, with the productivity slowdown leading into the Great Recession, transitions to the downturn state increased. The lower graph shows the posterior mode smoothed expected downturn risk facing agents. The downturn risk implied by the model estimates varies over time and matches the general history of the U.S. economy. Downturn risk reached high levels in the mid-1970s recession, driven by the OPEC oil embargo, later decreasing across the decade of the 1980s and 1990s. It played a small role in the two recessions of the early 1980s. Previous research (e.g. Bianchi (2013)) has found monetary policy to be especially important in causing these recessions. The U.S. went through an extended period of strong growth in the 1990s, and downturn risk is found to be quite low during this time period.

The decade of the 2000s had quite high downturn risk, reflecting the productivity slowdown in the middle of the decade and the recent financial crisis initiated by a large devaluation of the housing stock.

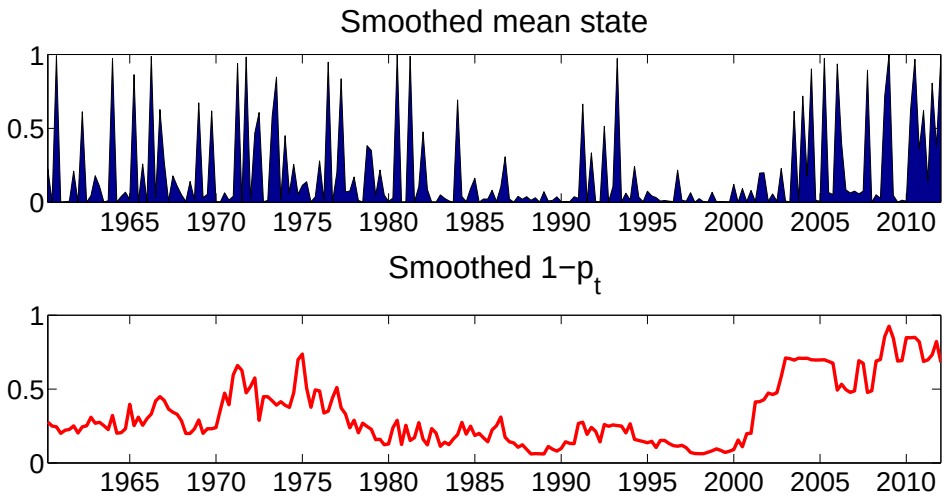


Figure 10 Posterior mode smoothed probability process. When the blue area is closer to 1, the model places a higher probability of being in the downturn state. Likewise, when the red line is closer to 1, there is a higher probability of a transition to the downturn state in the following period.

Figure 11 plots the untransformed $\tilde{p}_t$ process implied by the model against University of Michigan Index of Consumer Sentiments data. The survey asks questions involving consumer assessments of the current and future state of the economy and reflects their confidence. In the stylized facts section, I presented evidence that the fluctuations in this index respond more to expected downside movements in consumption growth as opposed to upside. Moreover, I also showed that changes in the index have important effects on the economy. Aside from the time period in the early 1980s dominated by the two recessions caused by tightening monetary policy, the $\tilde{p}_t$ process matches well with this sentiments measure. The correlation from the beginning of the sample to 1978 is 0.62 and the correlation after 1983 is 0.63. Hence, the model gives a new interpretation to fluctuations to consumer sentiments, namely that they respond to agents' changing assessments of downturn risk.

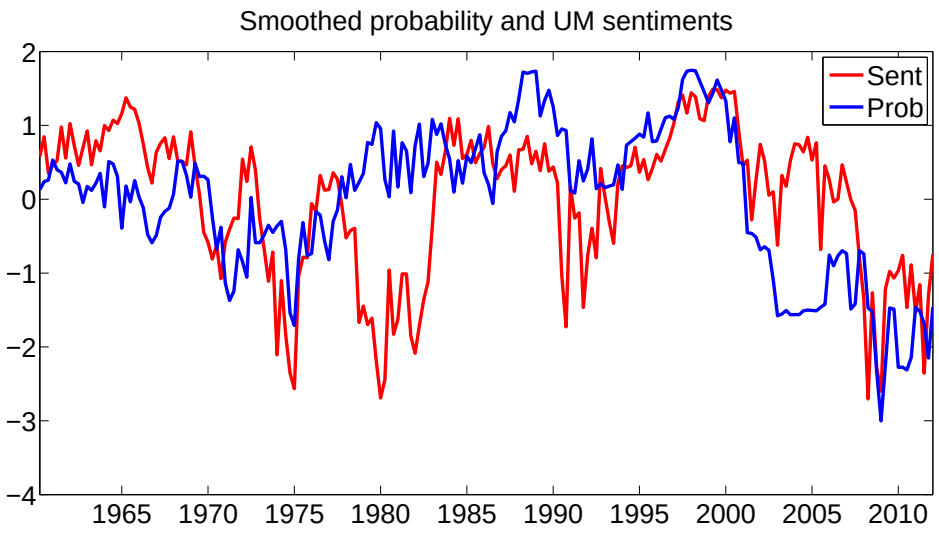


Figure 11 Posterior mode smoothed process $\tilde{p}_t$ and UM Index of Consumer Sentiments

Fluctuations and downturn risk shocks

It is interesting to analyze the historical importance of the downturn risk shock. To this end, I shut down all other shocks in the model and look at the amount of fluctuations attributed by the model to only downturn risk. The downturn risk shock is an important determinant of consumption growth and output growth. For consumption growth, it captures both the higher-frequency fluctuations and the medium-frequency movements. It matches many of the large declines during the two recessions of the 1970s and the large downward shift in the Great Recession. Changing downturn risk captures the rise of output growth in the late 1990s as well as the large decline during the Great Recession. Aside from these periods, it generally moves fluctuations in output growth at lower frequencies.

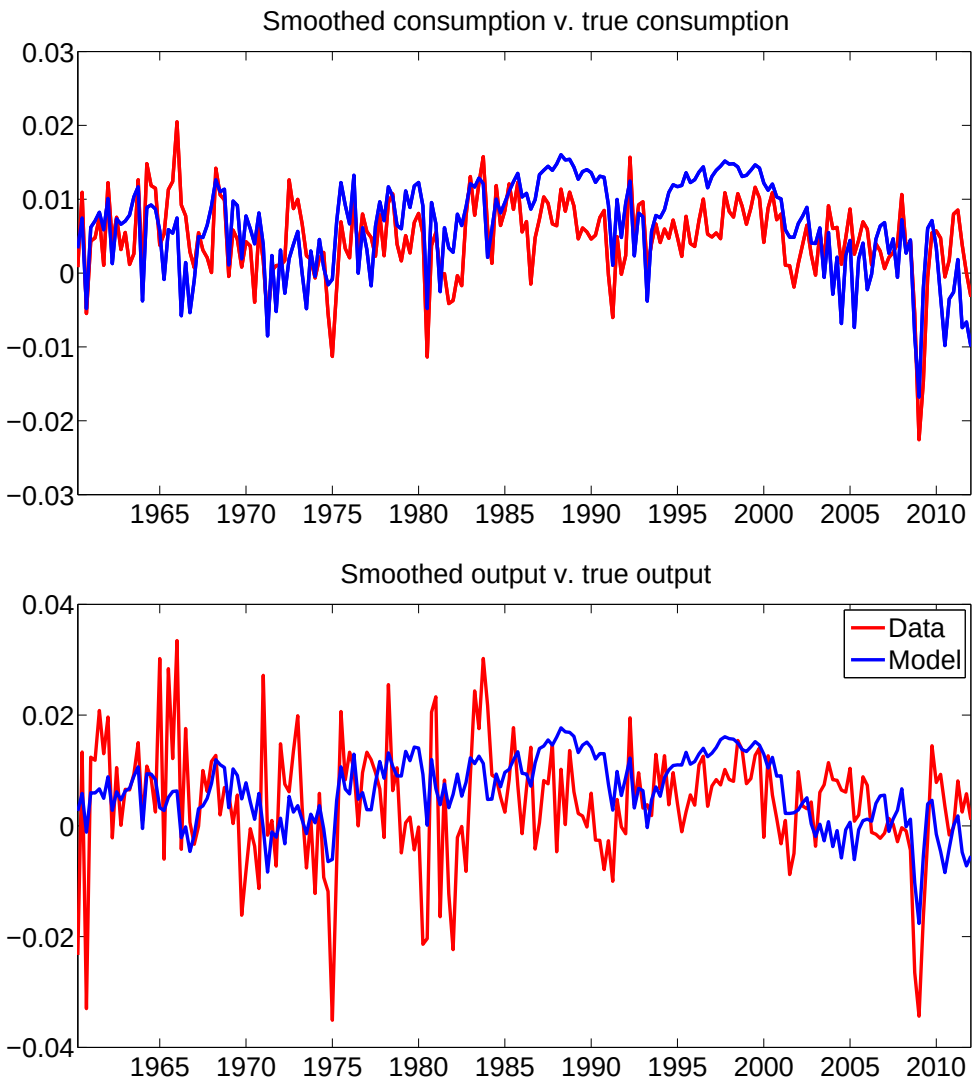


Figure 12 Posterior mode smoothed consumption and output growth implied by only downturn risk

Table 9 displays the variance decompositions at business cycle frequencies of several key real macroeconomic aggregates coming from the two competing expectations shocks in the model - the downturn risk shock and productivity growth news. The downturn risk shock drives a significant portion of output and hours fluctuations, along with being a dominant determinant of consumption fluctuations. This decomposition makes a lot of sense, as the downturn risk shock impacts longer-horizon fluctuations in the macroeconomy. Con-

	Downturn risk	Prod growth news
Out	30	3
Cons	69	10
Inv	7	1
Hours	25	4

Table 9 Posterior mode variance decomposition at business cycle frequencies (8−32 quarters) of real macroeconomic aggregates. Posterior intervals are in the appendix.

	Downturn risk	Prod growth news
Out	61	8
Cons	79	6
Inv	33	8
Hours	20	2

Table 10 Posterior mode variance decomposition at medium cycle frequencies (32 − 200 quarters) of real macroeconomic aggregates. Posterior intervals for downturn risk shock are in the appendix.

sumption allocations, being quite forward-looking, naturally respond to agents' expectations about the future. In comparison, productivity growth news, in line with findings by Schmitt-Grohe and Uribe (2012), Fujiwara, Hirose, and Shintani (2011), and Khan and Tsoukalas (2012), explain very little business cycle fluctuations. Moving to the medium-frequency cycle, the downturn risk shock becomes the dominant driver of macroeconomic fluctuations. It explains most fluctuations in output and consumption while gaining greatly in importance for investment. The shock continues to have an important impact on hours fluctuations, helping drive out the labor supply shock at lower frequencies. Therefore, in line with the historical probability estimates, the downturn risk shock captures the idea of decade-long stretches (such as the 1990s) of strong growth, infrequent recessions, and low downside risk, leading to strong consumer sentiments and decade-long stretches (such as the 1970s) of weak growth, frequent and deep recessions, and high downside risk, resulting in weak consumer sentiments. At the lower frequency, productivity growth news continues to have little impact on the key macroeconomic aggregates. Downturn transitions generate the negative skewness in key macroeconomic aggregates that are the hallmark of the asymmetric business cycle. Upon a transition to the downturn state, the large decrease in

productivity growth along with the large increase in depreciation causes agents to quickly decrease output, consumption, and investment. In the normal state, the model generates dynamics consistent with any standard macroeconomic model with Gaussian shocks.

	Data	Downturn risk shocks	Model
Out gr	-0.64	-0.84	-0.19
Cons gr	-1.00	-0.87	-0.60
Inv gr	-0.50	-0.74	-0.03

Table 11 Posterior mode model-implied skewness. Note that skewness of data differs from earlier tables because it is calculated only for $1960Q2 - 2011Q4$, the data on which the model was estimated.

The estimated model with only downturn risk shocks operative does a reasonable job of generating the skewness of the three series consistent with the data. It produces skewness values similar to what is seen in the data. Upon adding all of the other structural shocks, the baseline model still can account for 30% of the skewness found in output growth data and 60% of the skewness for consumption growth[20]. On the other hand, because the marginal efficiency of investment shock is the key driver of investment growth, the skewness in investment growth from the downturn risk shock does not carry through to the full model. The ability of the model to generate skewness in consumption growth is important, as recent research on consumption-based asset pricing has moved towards looking at skewness risk in consumption growth as an explanation for puzzles such as the equity risk premium. Having an estimated general equilibrium model that generates such skewness is a first step towards incorporating this feature in production-based asset pricing models.

Response of macroeconomy to downturn risk shocks

By altering agents' expectations about the future through changing the distribution of productivity growth and depreciation, movements in the probability process have important

[20]Posterior distributions for skewness can be found in the appendix.

effects on the macroeconomy. Figure 13 shows this response. In response to an increase in the risk of the downturn state, expected productivity slowly converges to a new permanently lower level. The slow diffusion of the productivity level is reminiscent of a trend productivity growth shock. This explains why the downturn risk shock explains so much of fluctuations at the medium frequency. Different from a pure trend productivity growth shock, the expected depreciation rate moves as well, peaking at around 30% above steady state in the first period after the innovation.

Consumption decreases immediately due to the wealth effect. Agents in the economy feel poorer from the expected negative movement in productivity, so they consume less in response. Investment also responds negatively after a negative probability shock. Two effects cause this comovement. First, the increase in expected future depreciation makes investment less attractive. Second, agents understand that the low productivity growth state will occur with a higher probability in the future, thereby generating future decreases in investment. The investment adjustment costs encourage agents to begin decreasing investment immediately to smooth out its expected movements. Consumption and investment reacting in the same direction naturally leads to output decreasing. Hours decrease because both consumption and investment decrease, thereby lowering demand of final goods and labor demand. Wages decline because of the lower firm demand for labor.

There exists two opposing forces moving inflation. First, the decreased production without lower productivity leads to a lower marginal cost for firms in the current period, which tends to generate lower inflation. Opposing this is the fact that marginal costs in the future are expected to increase, as low productivity levels increase marginal costs. With high enough price adjustment costs, firms immediately begin raising prices as they weight the future quite strongly. This effect dominates and inflation increases following a downturn risk shock. The federal funds rate decreases in response to a downturn risk shock. The Taylor rule reacts to increased inflation by increasing the nominal interest rate while it reacts to lower output growth by lowering the nominal rate. The effect of the output decrease dominates and the

federal funds rate goes down. It responds in a hump-shaped manner due to smoothing in the Taylor rule.

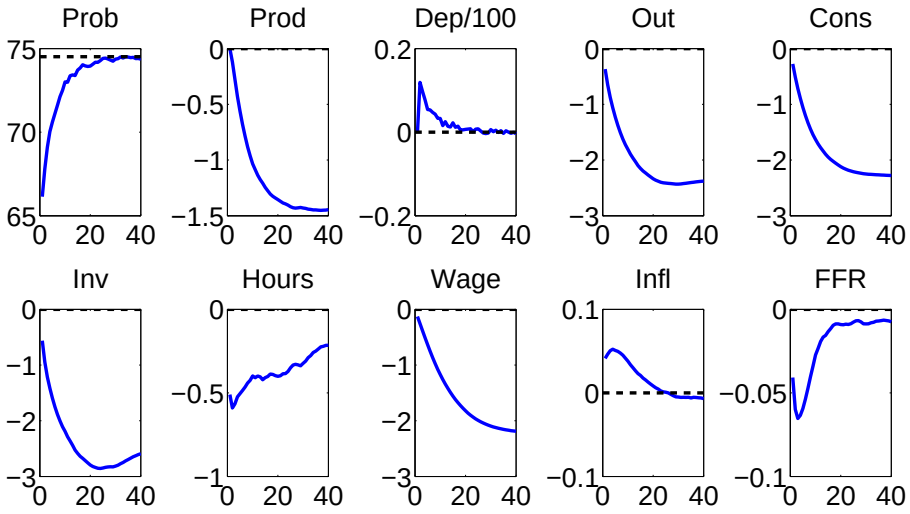


Figure 13 Posterior mode impulse response to downturn risk shock: probability change from $p_t = 0.74$ to $p_t = 0.66$ starting at steady state (multipled by 100 log deviations from steady state)

The response to a probability shock in figure 13 is consistent with the impulse response from an innovation to consumer sentiments in the VAR in figure 3. Namely, consumption, investment, and real wages decrease permanently, hours decrease and revert back to steady state quite slowly, inflation increases while the federal funds rate decreases in a hump-shaped manner.

Pure downturn risk shock

The pure downturn risk shock is important because it isolates the macroeconomic fluctuations from the expectations effect due to changing downturn risk. This shock generates changes in agents decisions with no movements in the fundamentals of the economy.

First, I look at the historical importance of the pure downturn risk shock. As figure 14 indicates, the shock drives the lower frequency trends in hours worked. It does well at cap-

turing the drop in labor supply in the early 1970s and the decline starting in the early 2000s. Dynamic equilibrium models in general have a difficult time matching hours. This has led researchers to add in a labor supply shock, which Justiniano, Primiceri, and Tambalotti (2010) find to drive the low-frequency movements. The finding that the pure downturn risk shock can capture the low frequency movements of hours renders the labor supply shock irrelevant for output growth, consumption growth, investment growth, and hours.

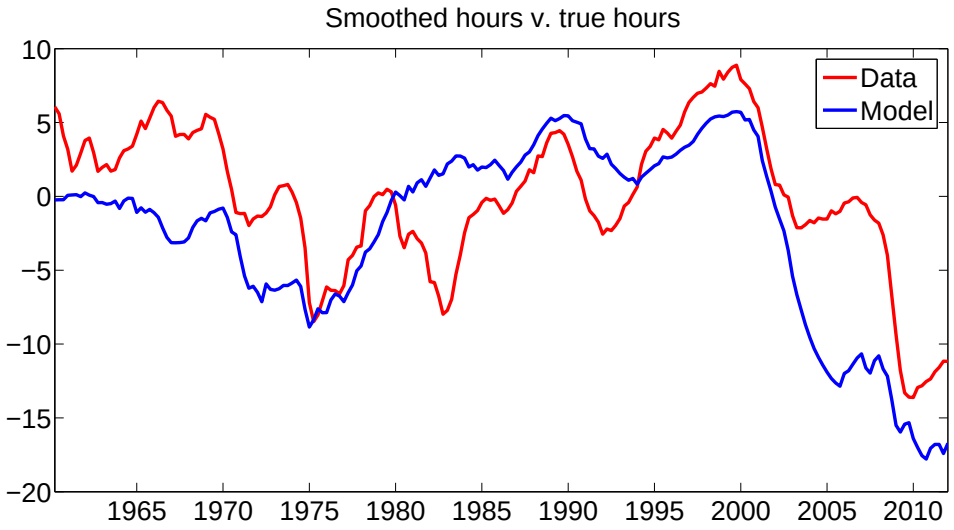


Figure 14 Posterior mode smoothed hours implied by only pure downturn risk

Turning to the impulse responses, the pure downturn risk shock generates transitory co-movement in consumption, investment, output, and labor, a defining feature of the business cycle.

The key macroeconomic aggregates respond on impact to a pure downturn risk shock in exactly the same way as to a downturn risk shock. Instead of productivity slowly reverting to a new long-run level and depreciation responding in a hump-shaped manner, however, productivity and depreciation do not move. As expected productivity and depreciation movements do not realize, the macroeconomic aggregates begin reverting back to their steady state values as the probability of the downturn state trends back to steady state. Inflation, originally increasing due to expectations of future productivity decreases, decreases

below 0 three periods following the downturn risk increase as firms have lower production with no corresponding decrease in productivity. With less upward inflationary pressures in the initial periods, monetary policy responds by more aggressively cutting interest rates.

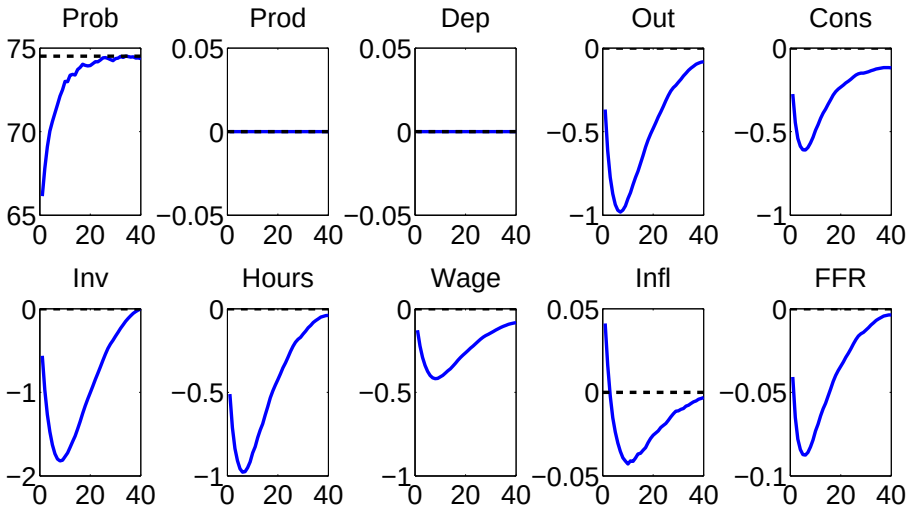


Figure 15 Posterior mode impulse response to pure downturn risk shock: probability change from $p_t = 0.74$ to $p_t = 0.66$ starting at steady state (multipled by 100 log deviations from steady state)

High price and wage adjustment costs help the model generate large movements in labor supply from the pure downturn risk shock. A pure downturn risk shock lowers agents' expectations about future productivity growth and raises agents' expectations about future depreciation. As discussed, the interplay between these two effects lowers consumption and investment. Price rigidities encourage firms to increase prices immediately, as they expect higher future marginal costs from the expected decreases in productivity. The high prices further lowers demand for firm goods, amplifying the decrease in output, consumption, and investment. As firms do not have to produce as much, labor demand is low. Wage rigidities prevent real wages from falling as much upon impact, which further lowers demand for labor, again amplifying the downturn.

Figure 16 shows the responses of the economy first shutting down price rigidity (red lines)

and then shutting down wage rigidity (magenta lines). Without price rigidites, inflation drops immediately due to the lower marginal cost from less production without changes in productivity. Low prices increase demand for the firm goods, which dampens the downturn. The low inflation, however, also increases real wages, which tends to lower labor demand. Overall, the first effect dominates, and the declines in output, consumption, investment, and hours are more mild. Shutting down wage rigidities causes real wages to drastically decrease, which greatly limits the decline in hours. In turn, this allows for more production and a milder decline in the other real aggregates. Therefore, both price and wage rigidities are critical to generate a large downturn from the expectations effect and allow the pure downturn risk shock to drive hours movements.

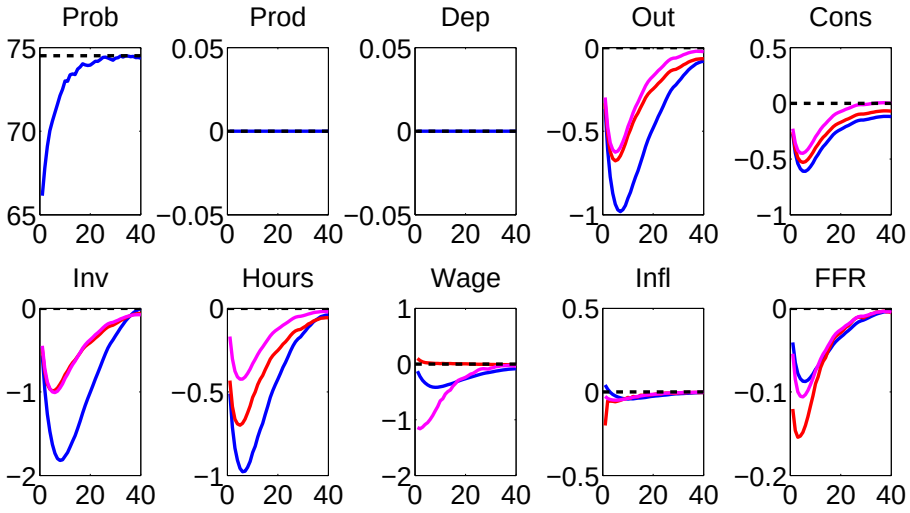


Figure 16 Posterior mode impulse response to pure downturn risk shock: probability change from $p_t = 0.74$ to $p_t = 0.66$ starting at steady state (multipled by 100 log deviations from steady state, Blue: Baseline, Red: No Price, Magenta: No Wage)

	Business cycle freq $(8-32$ qtrs)	Medium cycle freq $(32-200$ qtrs)
Out	27	11
Cons	33	5
Inv	11	18
Hours	31	44

Table 12 Posterior mode variance decomposition of real macroeconomic aggregates to pure downturn risk shock. Posterior intervals are in the appendix.

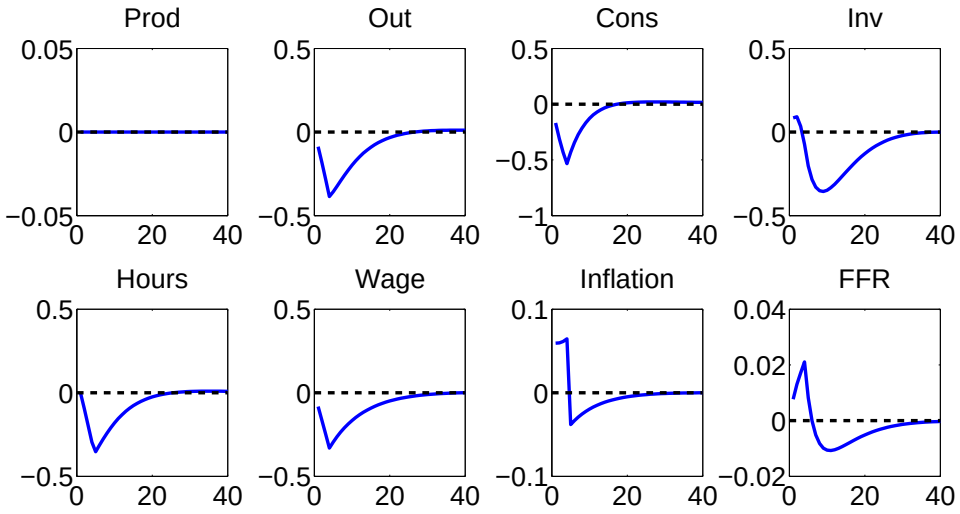


Figure 17 Posterior mode -1 standard deviation unrealized productivity growth news shock (multiplied by 100 log deviations from steady state)

Figure 17 shows the economy's response to -1 standard deviation unrealized productivity growth news shock at the posterior mode parameters. In contrast to the pure downturn risk shock, the unrealized news shock does not generate comovement between consumption and investment upon impact. Consumption decreases from the wealth effect. Because the nominal rigidities and investment adjustment costs are not large enough, investment increases. Moreover, notice that in response to productivity growth news, the federal funds rate actually increases, which is also inconsistent with the VAR evidence in figure 3. Table 12 shows the importance of the pure downturn risk shock at business and medium cycle frequencies. At business cycle frequencies, the expectations effect alone is an important driver of fluctuations, explaining around 1/3 of fluctuations in output and consumption. The

pure downturn risk shock also drives 1/3 of hours movements. Therefore, at the estimated parameters, the model attributes a significant portion of business cycle fluctuations to the changing expectations of agents. Moving to the medium-frequency cycle, the model generates a significant portion of hours fluctuations through the expectations channel. In fact, the introduction of this pure downturn risk effect is the single largest factor driving out the labor supply shock from longer-horizon frequencies. Therefore, the model gives a new interpretation to the low-frequency movements of hours, namely that changing agents' assessments about the downturn risk in the economy interpretable as sentiments drive the longer-horizon movements. Figure 18 plots quarterly hours and consumer sentiments data. The series comove quite closely, especially at lower frequencies, with a correlation of 0.63. The structural VAR in the stylized facts section suggests that between 38 and 73 percent of per capita hours variance (80% posterior interval) 40-quarters out comes from Cholesky-identified innovations to the ICS[21].

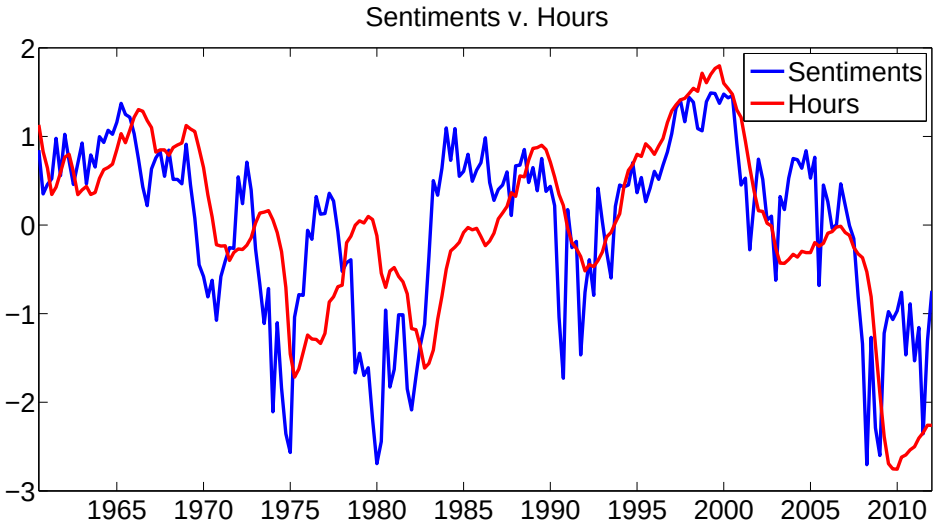


Figure 18 Hours and the Index of Consumer Sentiments 1960 – 2012

[21] Even with the ICS ordered last, the SVAR attributes between 13 and 49 percent of hours fluctuations from ICS innovations (80% posterior interval).

Transition to the downturn state

A transition to the downturn state generates a large permanent decrease in productivity and transitory increase in depreciation. This leads to permanent declines in output, consumption, investment, and real wages much larger in magnitude than an unanticipated productivity growth shock. These transitions are key for the model to generate the skewness seen in the data.

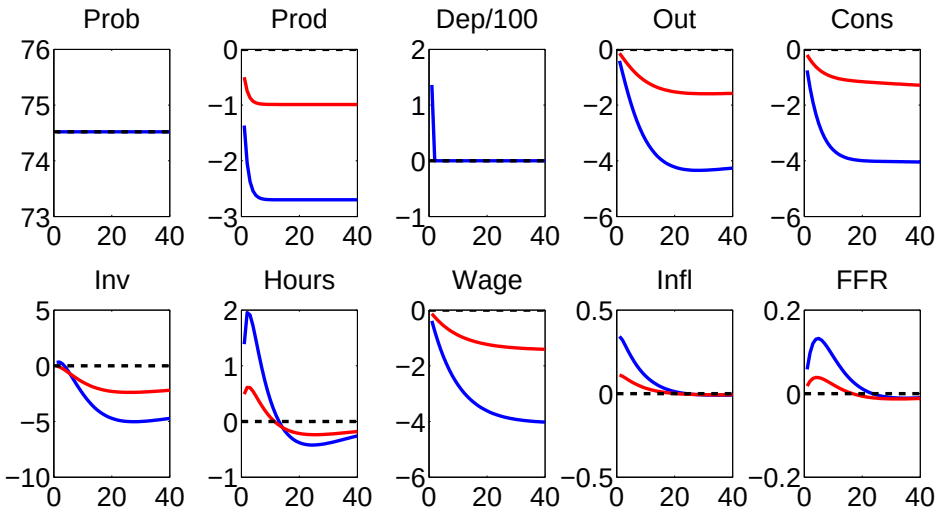


Figure 19 Posterior mode impulse response to a realization of the downturn state starting at the steady state (blue) compared to -1 standard deviation productivity growth unanticipated shock (red) (multipled by 100 log deviations from steady state)

Figure 19 shows the response of the economy to a downturn state transition in blue and the response to a -1 standard deviation productivity growth innovation in red. Hours increase greatly. This occurs because the decrease in productivity combined with nominal price rigidity increases the demand for labor. The standard intuition carries through from the response of hours to an unanticipated productivity growth shock, but is amplified because of the large decline in productivity growth from a downturn state transition. Hours become more hump-shaped relative to the response after a productivity growth decline because the corresponding rise in depreciation lowers the initial increase in hours. Inflation increases because of the increase in current and future marginal costs from the decrease in productivity

growth. Again, the federal funds rate responds in opposite directions to inflation and output growth as specified by the Taylor rule. At the posterior mode parameters, the response is positive coming from the strong inflation increase.

Investment slightly increases on impact, which may be undesirable, although the strong decrease in the following periods still allows the model to generate negative skewness in investment growth from a downturn transition. I also estimated a version of the model allowing for an autoregressive parameter on the depreciation shock. Doing so allows for persistence in depreciation movements, which reduces the incentive to invest upon a downturn risk shock. This addition makes investment decrease on impact without affecting other results[22].

Downturn risk shocks and the Great Recession

The recent Great Recession was one of the deepest over the past few decades. Many think that the recession marked the end of the Great Moderation, a time of low volatility of many macroeconomic and financial series. The rapidly changing macroeconomic conditions could lead to agents' expectations or sentiments having a large impact on the macroeconomy.

[22]Results available upon request.

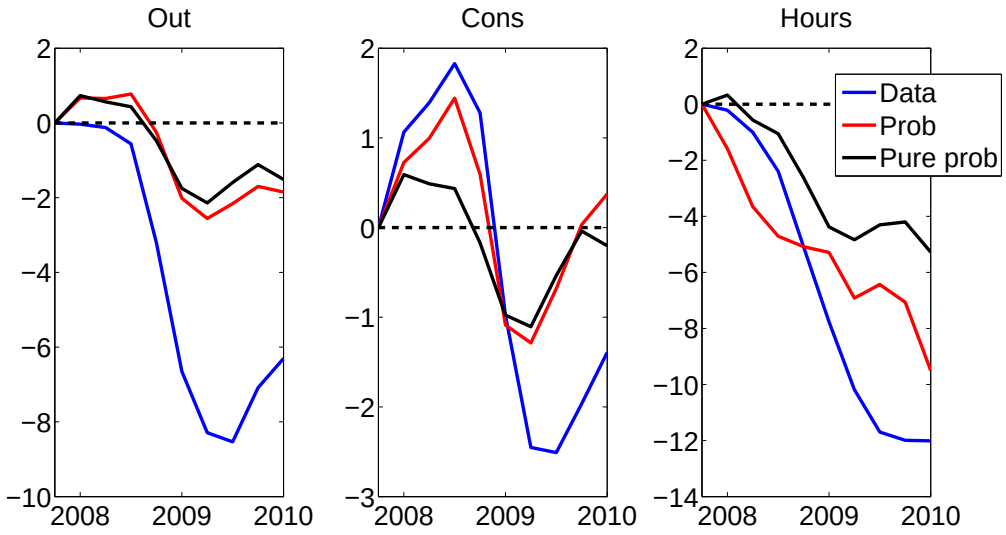


Figure 20 Per capita output, consumption, and hours during the Great Recession from data and implied by the model with only downturn risk shocks. Multiplied by 100. Value at first quarter of recession as determined by the NBER (2007$Q4$) normalized to 0. Graphs run until 2009$Q4$, 2 quarters after end of recession as determined by the NBER (2009$Q2$).

It therefore is interesting to examine the importance of downturn risk shocks in driving the dynamics of several key macroeconomic variables, shown in figure 20, across the recession. Both the pure downturn risk and downturn risk shocks lead to dynamics of the three series that track the observed fluctuations closely. Expectations of agents alone, captured by the black line in the figures, were responsible for nearly a quarter of the output decline, 40% of the consumption decline, and 40% of the hours decline from the beginning of the recession until 2009$Q1$. Assessments of downturn risk by agents decreased from 2009$Q1$ – 2009$Q2$, portending the exit from the recession. Therefore, the model suggests that heightened downside risk may have been an important contributing factor to the depth of the recession.

Removing downturn risk shocks

The importance of downturn risk shocks can be seen by how much the implications of the model change upon removing them. Table 13 gives model comparison results to a variety of alternative specifications. In a specification without downturn risk shocks, the marginal

Model	Log Marginal Likelihood
Downturn risk shock	5157
Trend prod	−23
Two-state Markov-switching	−47
MEI news	−52
No downturn risk shock	−185

Table 13 Log marginal likelihood comparison

	Without downturn risk shocks	With downturn risk shocks
Out	24	3
Cons	38	10
Inv	9	1
Hours	11	4

Table 14 Posterior mode variance decomposition of real macroeconomic aggregates to productivity news shock at business cycle frequencies ($8-32$ quarters). Posterior intervals are in the appendix.

likelihood drops by a significant amount, over 180 log marginal likelihood points[23]. The implications of the model change drastically as well. Table 14 compares the importance of productivity growth news at business cycle frequencies in the model without downturn risk shocks and with downturn risk shocks. Without downturn risk shocks, productivity growth news becomes an important determinant of macroeconomic fluctuations, especially for consumption and output. Upon the addition of downturn risk shocks, the importance of productivity growth news becomes vastly diminished. As the analysis in the simple model suggests, this comes from the downturn risk shock's ability to more easily generate comovement between consumption and investment. Table 15 contrasts the importance of the labor supply shock with and without the downturn risk shock. Without the downturn risk shock, the labor supply shock's importance for hours is comparable to the results of Justiniano, Primiceri, and Tambalotti (2010). Specifically, it has moderate impact at business cycle frequencies but increases greatly in importance at longer horizons. The introduction of the downturn risk shock renders the labor supply shock irrelevant for the four key aggregates at business cycle frequencies and drastically decreases its importance at longer horizons as well. The labor supply shock remains important for real wage fluctuations, of which it

[23]Posterior predictive checks can be found in the appendix.

Business cycle freq $(8-32$ qtrs)		
	Without downturn risk shocks	With downturn risk shocks
Out	4	2
Cons	4	1
Inv	2	1
Hours	18	4
Medium cycle freq $(32-200$ qtrs)		
	Without downturn risk shocks	With downturn risk shocks
Out	4	2
Cons	3	1
Inv	4	5
Hours	61	14

Table 15 Posterior mode variance decomposition of real macroeconomic aggregates to labor supply shock. Posterior intervals are in the appendix.

determines 53% of fluctuations at business cycle frequencies. To illustrate that the gains in model fit from the downturn risk shock are not trivial, I also estimate the model removing the downturn risk shock and including first MEI news, then trend productivity growth, and finally two-state Markov-switching on the downturn state. All fit the data worse than the model with the downturn risk shock. The model with MEI news captures a scenario with news in both productivity and the capital accumulation equation. It shows that much of the gains from the downturn risk shock come from the interaction between productivity growth and the depreciation rate, not the fact that there exists news in both components. The model with trend productivity growth has a long-run risks news shock that has implications at the medium frequency. That this framework fits the data worse illustrates the importance of both having depreciation news and linking the medium-frequency cycle to macroeconomic asymmetries. Finally, the two-state Markov-switching model does not have the element of pure changes in risk of the downturn state. The fact that this model fits the data worse illustrates that it is not enough to just have a downturn state. Instead, it really is the changes in the risk of this downturn state that drive fluctuations. Finally, I ask the question whether any of the expectations shocks in the competing models can generate the comovement we see from the impulse response to a -1 standard deviation innovation to the Index of Consumer Sentiments. For reference, the first column of table 16

Variable	VAR	Downturn risk	Prod news	MEI news	Trend prod
Cons	−	−	−	+	−
Inv	−	−	+	−	+
Hours	−	−	−	+	+
Wage	−	−	−	+	−
Infl	+	+	+	−	+
FFR	−	−	+	−	+

Table 16 Comovements of the Stylized Facts 4 SVAR impulse response to −1 standard deviation innovation to the Index of Consumer Sentiments versus posterior mode comovements of endogenous variables produced by a variety of expectations shocks

reproduces the comovements in the SVAR from Stylized Facts 4. The second column shows the comovements implied by the downturn risk shock, which are again consistent with the SVAR results. As discussed earlier, the productivity growth news shock does not generate this comovement. Table 16 illustrates that the MEI news shock and the trend productivity growth shock also do not match these comovements.

1.5. Conclusion

In this paper, I introduce variable risks of downside shocks into a dynamic equilibrium model. I show that this modification can capture four key stylized facts of business cycles difficult for standard real business cycle or New Keynesian models to generate: the negative skewness of output, consumption, and investment growth, the changing frequencies and depths of recessions, the relationship between consumer sentiments and future downside consumption growth innovations, and the comovements in macroeconomic series generated by innovations in consumer sentiments.

I first give evidence in support of the stylized facts. This evidence provides guidance as to what elements of the data the model should capture. Crucially, having a downturn realization simultaneously decrease productivity growth and increase depreciation allows the model to produce the stylized facts.

The downturn risk shock is important at business cycle frequencies and a dominant compo-

nent of the medium-frequency cycle, driving 60% of output fluctuations at lower frequencies. The pure downturn risk shock, which captures the effect of agents' rational expectations over the amount of downturn risk in the economy, is an important driver of business cycles and a key component of hours movements at lower frequencies. Moreover, responses of the economy to a change in the downturn risk are consistent with the VAR response from an identified innovation to the University of Michigan Index of Consumer Sentiments. Critically, an increase in downturn risk generates declines in consumption, investment, and output, which is the crucial comovement that characterizes business cycles.

I believe two future avenues of research prove interesting. First on the agenda is extending the model to include asset prices. Research by Gourio (2012) has shown that the disasters mechanism can simultaneously drive business cycles and capture various properties of asset prices such as the countercyclicality of risk premia. Formally empirically evaluating the ability of the less extreme downturn risks at capturing asset price movements is therefore a natural next step. Moreover, the forward-looking nature of asset prices can provide discipline on agents' expectations, helping to sharper identify the amount of downside risk currently in the economy. Performing the estimation will require a model solution beyond the first order, which is quite a new concept for Markov-switching dynamic equilibrium models, but this should be feasible given the recent theoretical advances of Foerster, Rubio-Ramirez, Waggoner, and Zha (2013).

Second, it is important to further investigate the mechanisms generating the downturn risks. Exploring a deeper structural model that can endogenously lead to fluctuations in the risk of a deep downturn state could prove enlightening. Given the reduced-form evidence on the link between downside shock predictability and measures of financial systemic risk (Giglio, Kelly, and Pruitt (2013)), an explicit model involving fragility in the financial sector interfering with bank intermediation would be the first avenue to explore.

1.6. Appendix

1.6.1. Data

All data series are quarterly from $1953Q1 - 2011Q4$. I start the structural estimation in 1960 because that is the first date quarterly University of Michigan Index of Consumer Sentiments series is available.

I use civilian noninstitutional population as my measure of population. I use seasonally adjusted gross domestic product implicit price deflator as the data on the price level in the economy.

I define nominal consumption as seasonally adjusted personal consumption expenditures on nondurable goods plus seasonally adjusted personal consumption expenditures on services. Nominal investment is seasonally adjusted gross domestic private investment plus seasonally adjusted personal consumption expenditures on durable goods. Nominal government consumption is seasonally adjusted state and local consumption expenditures, state and local government gross investment, federal nondefense consumption expenditures, federal nondefense gross investment, and federal national defense consumption expenditures and gross investment. Nominal output is defined as nominal consumption plus nominal investment plus nominal government consumption. Nominal wage is compensation per hour in the nonfarm business sector. To convert to real, per capita terms, all series are divided by civilian noninstitutional population and gross domestic product implicit price deflator.

Inflation is defined as the log difference of gross domestic product implicit price deflator. I use the Effective federal funds rate as my measure for the federal funds rate. Hours worked data is from the seasonally adjusted hours of all persons index from the business sector divided by civilian noninstitutional population.

The University of Michigan Index of Consumer Sentiments comes from the Thomson Reuters/

Quantile	Relative Loss (GDP defl)	CPI defl less food and energy	PCE defl
10	0.91*	0.91*	0.88*
20	0.90**	0.93*	0.92*
50	0.96	0.97	0.94**
80	1.00	1.05	0.97
90	1.12	1.02	1.01

Table 17 Results for time $t+1$ consumption growth innovation given time t log Index of Consumer Sentiments. Using one-sided tests of higher predictive power of the sentiments, $**$ indicates significance at the 5% level, and $*$ indicates significance at the 10% level.

University of Michigan Survey of Consumers.

1.6.2. Stylized Fact 3 Quantile Regression

The quantile loss function is defined in equation 1.33.

$$\rho_\tau(x) = x\,(\tau - I_{x<0}) \tag{1.33}$$

Quantile regression as in Koenker and Bassett (1978) seeks to minimize the expected quantile loss for given τth quantile by choosing parameters α and β in equation 1.1[24].

To demonstrate the robustness of the results, I consider alternative deflators of nominal consumption as well as 10% and 90% quantiles of the consumption growth innovations. Table 17 presents these results. In all cases, the ICS forecasts the lower quantiles significantly better than the historical quantile. The forecasting gains decrease upon considering higher quantiles. At the median, only the PCE deflated consumption growth innovations show evidence of predictability with the ICS. None of the higher quantiles show any evidence of predictability, oftentimes with the model including the ICS forecasting worse than the historical quantile.

[24]To perform the quantile regression, I use the m-file "rq_fn" provided by Roger Koenker.

1.6.3. Stylized Fact 4 Bayesian $VAR(4)$

I run a Bayesian $VAR(4)$ on quarterly log University of Michigan Index of Consumer Sentiments, per capita log consumption, per capita log investment, per capita log hours, per capita log real wages, inflation, federal funds rate, in that order. I choose lag length 4 to be in line with the specification of Barsky and Sims (2012). Equation 1.34 gives the reduced-form VAR. I use a Cholesky identification to identify an innovation to the ICS controlling for innovations to the other series in the economy.

$$X_t = c + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \Phi_3 X_{t-3} + \Phi_4 X_{t-4} + \epsilon_t, \epsilon_t \sim N(0, \Sigma) \tag{1.34}$$

I put a Normal-Inverse Wishart prior on the parameters. The Inverse Wishart prior on Σ has degrees of freedom parameter 8 (2 more than the dimension of the data and the minimum amount for which the prior mean exists) and the diagonal elements of the scale matrix as the variance of the OLS residuals from an $AR(1)$ for each variable. For the prior on $\{c, \Phi_i\}$, I use the Minnesota prior (Litterman (1979)) as laid out in Giannone, Lenza, and Primiceri (2012). Following Giannone, Lenza, and Primiceri (2012), I draw from the posterior of a hyperparameter governing the tightness of the prior covariance between the regressors.

I draw 20000 times from the posterior and construct pointwise 68% credible sets for the impulse response functions. To form the impulse responses, I take a draw from the posterior and then perform a Cholesky-identified impulse response to the ICS variable[25].

1.6.4. Log-linearized equilibrium conditions

Call $z_t = A_t^{\frac{1}{1-\alpha}}$.

Define $\tilde{mc}_t = \left(\frac{1}{1-\alpha}\right)^{1-\alpha} \left(\frac{1}{\alpha}\right)^{\alpha} \frac{w_t^{1-\alpha} r_t^{\alpha}}{A_t}$

[25]I use code provided by Giannone, Lenza, and Primiceri (2012)

Stationarized variables are $\tilde{c}_t = \frac{c_t}{z_t}, \tilde{i}_t = \frac{i_t}{z_t}, \tilde{k}_{t+1} = \frac{k_{t+1}}{z_t}, \tilde{\lambda}_t = z_t\lambda_t, \tilde{\zeta}_t = z_t\zeta_t, \tilde{w}_t = \frac{w_t}{z_t}, \tilde{z}_t = \frac{z_t}{z_{t-1}}$.

Define $\hat{x}_t = \log x_t - \log x_{ss}$.

Define Q as the transition matrix between the states $\tilde{s}_t$. Expectations $E_{\tilde{s}_t,t}$ depend linearly on the current state $\tilde{s}_t$.

$$\tilde{\lambda} = \frac{1}{\tilde{c}}\left(1 - he^{-\Lambda_z}\right)^{-1} - h\beta\frac{1}{\tilde{c}}\left(e^{\Lambda_z} - h\right)^{-1}$$

$$\Pi = \beta e^{-\Lambda_z} R$$

$$r = \Phi_1$$

$$\tilde{\zeta} = \beta e^{-\Lambda_z}\left(\tilde{\lambda}r + \tilde{\zeta}(1-\delta)\right)$$

$$\tilde{\lambda} = \tilde{\zeta}$$

$$\eta\psi l^{\gamma} = \tilde{\lambda}(\eta - 1)\tilde{w}$$

$$\tilde{w} = \frac{1-\alpha}{\alpha}\frac{\tilde{k}}{l}re^{-\Lambda_z}$$

$$\frac{1}{\epsilon} = 1 - \left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\tilde{w}^{1-\alpha}r^{\alpha}$$

$$\tilde{c} + \tilde{i} = \frac{1}{\tilde{g}}\tilde{y}$$

$$\tilde{y} = e^{-\Lambda_z + \Lambda_A}\tilde{k}^{\alpha}l^{1-\alpha} - \phi$$

$$\tilde{k} = (1-\delta)e^{-\Lambda_z}\tilde{k} + \tilde{i}$$

Finding the steady state

$$\tilde{\Pi} = \Pi$$

$$u = 1$$

$$\tilde{R} = \frac{\Pi}{\beta} e^{\Lambda_z}$$

$$\tilde{r} = \frac{1}{\beta} e^{\Lambda_z} - (1 - \delta)$$

$$\tilde{w} = \left(\frac{\left(1 - \frac{1}{\epsilon}\right)}{\left(\frac{1}{1-\alpha}\right)^{1-\alpha} \left(\frac{1}{\alpha}\right)^{\alpha} \tilde{r}^{\alpha}} \right)^{\frac{1}{1-\alpha}}$$

Solve for l using nonlinear equation solver on:

$$\eta \psi l^{\gamma} = \tilde{\lambda}(l)(\eta - 1)\tilde{w}$$

$$\tilde{\lambda}(l) = \frac{1}{\tilde{c}(l)} \left(1 - h e^{-\Lambda_z}\right)^{-1} - h\beta \frac{1}{\tilde{c}(l)} \left(e^{\Lambda_z} - h\right)^{-1}$$

$$\tilde{c}(l) = \frac{1}{\tilde{g}} \tilde{y}(l) - \tilde{i}(l)$$

$$\tilde{i}(l) = \left(1 - (1 - \delta) e^{-\Lambda_z}\right) \tilde{k}(l)$$

$$\tilde{y}(l) = e^{-\Lambda_z + \Lambda_A} (\tilde{k}(l))^{\alpha} (l)^{1-\alpha} - \phi$$

$$\tilde{k}(l) = \frac{\tilde{w}}{\tilde{r}} \frac{\alpha}{1 - \alpha} l e^{\Lambda_z}$$

Once l is found, all of the other equations are simply functions of l.

Log-linearized equilibrium conditions

$$(1-h\beta e^{-\Lambda_z})\hat{\tilde{\lambda}}_t = -\frac{1+h^2\beta e^{-2\lambda_z}}{1-he^{-\Lambda_z}}\hat{\tilde{c}}_t + \frac{h}{e^{\Lambda_z}-h}\hat{\tilde{c}}_{t-1} + \frac{\beta h e^{-\Lambda_z}}{(1-he^{-\Lambda_z})}E_t\left(\hat{\tilde{c}}_{t+1}+\hat{\tilde{z}}_{t+1}\right) - \frac{h}{e^{\Lambda_z}-h}\hat{\tilde{z}}_t$$

$$\hat{\tilde{\lambda}}_t = E_t\left(-\hat{\tilde{z}}_{t+1} + \hat{\tilde{\lambda}}_{t+1} + \hat{\tilde{R}}_t - \hat{\Pi}_{t+1}\right)$$

$$\hat{\tilde{r}}_t = \left(\hat{\tilde{\zeta}}_t - \hat{\tilde{\lambda}}_t\right) + \frac{\Phi_2}{\Phi_1}\hat{u}_t$$

$$\hat{\tilde{\zeta}}_t = \beta E_t\left(e^{-\Lambda_z}\left((r+1-\delta)\left(-\hat{\tilde{z}}_{t+1}\right) + \tilde{r}\left(\hat{\tilde{\lambda}}_{t+1}+\hat{\tilde{r}}_{t+1}+\hat{u}_{t+1}\right) + \hat{\tilde{\zeta}}_{t+1}(1-\delta) - \Phi_1\hat{u}_{t+1} - \delta\hat{\delta}_{t+1}\right)\right)$$

$$\hat{\tilde{\lambda}}_t = \hat{\tilde{\zeta}}_t - S''\left(e^{\Lambda_z}\right)e^{2\Lambda_z}\left(\hat{\tilde{z}}_t + \Delta\hat{\tilde{i}}_t\right) + \hat{\mu}_t + \beta E_t\left(S''\left(e^{\Lambda_z}\right)e^{2\Lambda_z}\left(\Delta\hat{\tilde{i}}_{t+1} + \hat{\tilde{z}}_{t+1}\right)\right)$$

$$\eta\psi(1+\gamma)\left(l^d\right)^{1+\gamma}\hat{l}_t^d + \eta\psi\left(l^d\right)^{1+\gamma}\hat{\psi}_t - \lambda\phi_w e^{2\Lambda_z}\tilde{y}_d\left(\hat{\tilde{z}}_t + \Delta\hat{\tilde{w}}_t + \hat{\Pi}_t - \alpha\hat{\Pi}_{t-1} - (1-\alpha)\hat{\Pi}_t^*\right) + ...$$

$$\beta E_t\left(\lambda\phi_w e^{2\Lambda_z}\tilde{y}^d\left(\Delta\hat{\tilde{w}}_{t+1} + \hat{\Pi}_{t+1} - \alpha\hat{\Pi}_t - (1-\alpha)\hat{\Pi}_{t+1}^* + \hat{\tilde{z}}_{t+1}\right)\right) = \lambda(\eta-1)\tilde{w}l^d\left(\hat{\tilde{\lambda}}_t + \hat{\tilde{w}}_t + \hat{l}_t^d\right)$$

$$\hat{\tilde{w}}_t = \hat{\tilde{u}}_t + \hat{\tilde{k}}_t - \hat{l}_t^d + \hat{\tilde{r}}_t - \hat{\tilde{z}}_t$$

$$\hat{\tilde{\lambda}}_t + \hat{\tilde{y}}_t^d + \phi_p\left(-\hat{\Pi}_t + \alpha\hat{\Pi}_{t-1} + (1-\alpha)\hat{\Pi}_t^*\right) + \beta E_t\left(\phi_p\left(\hat{\tilde{\Pi}}_{t+1} - \alpha\hat{\Pi}_t - (1-\alpha)\hat{\Pi}_{t+1}^*\right)\right) =$$

$$\epsilon(1-\tilde{mc})(\hat{\tilde{\lambda}}_t + \hat{\tilde{y}}_t^d) - \epsilon\tilde{mc}\hat{\tilde{mc}}_t$$

$$\hat{R}_t = \gamma_R\hat{R}_{t-1} + (1-\gamma_R)\left(\gamma_\pi\left(\hat{\Pi}_t - \hat{\Pi}_t^*\right) + \gamma_y\left(\hat{\tilde{z}}_t + \Delta\hat{\tilde{y}}_t^d\right)\right) + \epsilon_{r,t}$$

$$\tilde{c}\hat{\tilde{c}}_t + \tilde{i}\hat{\tilde{i}}_t = \frac{\tilde{y}^d}{e^{\tilde{g}}}\left(\hat{\tilde{y}}_t^d - \hat{\tilde{g}}_t\right)$$

$$\tilde{y}^d\hat{\tilde{y}}_t^d = e^{-\Lambda_z+\Lambda_A}(u\tilde{k})^\alpha(l^d)^{1-\alpha}\left(-\hat{\tilde{z}}_t + \hat{\tilde{g}}_{A,t} + \alpha\left(\hat{u}_t + \hat{\tilde{k}}_t\right) + (1-\alpha)\hat{l}_t^d\right)$$

$$e^{\Lambda_z}\hat{\tilde{k}}_{t+1} = \left(e^{\Lambda_z} - (1-\delta)\right)\left(\hat{\tilde{i}}_t + \hat{\mu}_t\right) + (1-\delta)(\hat{\tilde{k}}_t - \hat{\tilde{z}}_t) - \Phi_1\hat{u}_t - \delta\hat{\delta}_t$$

$$\hat{g}_{A,t} = \rho_A\hat{g}_{A,t-1} + c_A(\tilde{s}_t) + \epsilon_{A,t} + \epsilon_{4,A,t-4}$$

$$\hat{\delta}_t = d(\tilde{s}_t)$$

$$\hat{\mu}_t = \rho_\mu\hat{\mu}_{t-1} + \epsilon_{\mu,t}$$

$$\hat{\Pi}_t^* = \rho_\Pi\hat{\Pi}_{t-1}^* + \epsilon_{\Pi,t}$$

$$\hat{\tilde{g}}_t = \rho_g\hat{\tilde{g}}_{t-1} + \epsilon_{g,t}$$

$$\hat{\psi}_t = \rho_\psi\hat{\psi}_{t-1} + \epsilon_{\psi,t}$$

$$\tilde{s}_t = Q\tilde{s}_{t-1} + e_{\tilde{s},t}$$

1.6.5. Calculating impulse response to a probability shock

I use the following algorithm to calculate an length T to a probability shock.

Economy starts in steady state at $t = -1$.

1. Set two different states $\tilde{s}_{1,0}$ and $\tilde{s}_{2,0}$ at $t = 0$ representing two different economies.
2. Draw a uniform random variable.
3. Transition to the state at time period $t = 1$ in accordance to the transition probability of the state $\tilde{s}_{1,0}$ and $\tilde{s}_{2,0}$, respectively, for the two economies. Use the same uniform random variable to determine the state transition.
4. Calculate the difference in the variables of interest between the two economies.
5. Continue forward until time $t = T$.
6. Repeat above steps for N simulations.
7. Take an average over the N simulations of the difference in the variables of interest.

Unrealized probability shock

1. Save the implied paths of productivity growth $\hat{g}_{A,t}(\tilde{s}_t)$ and depreciation rate $\hat{\delta}_t(\tilde{s}_t)$ from the impulse response.
2. Extract the implied innovations of productivity growth and depreciation rate
3. Rerun the impulse response, this time with unanticipated productivity growth and depreciation rate innovations to cancel out the realizations of the two variables.

1.6.6. Estimation on simulated data

As advocated by Schmitt-Grohe and Uribe (2012) for a check of identification, I simulate data of length 208 from the model and draw from the posterior distribution. I impose flat priors on the parameters to avoid masking identification issues. I focus on any potential identification problems in the parameters involving the downturn risk shock process (such as productivity growth shock parameters and MEI shock parameters) and fix the other parameters at their true values.

Therefore, I take 10000 draws from the posterior with 5000 burn-in from the parameters of the technology shock, MEI shock, and downturn risk shock. As can be seen in table 18, for the sample size, the estimation performance is reasonable and there does not seem to be any identification problems.

Parameter	True	Lower (10%)	Mean	Upper (90%)
ρ_A	0.39	0.38	0.41	0.44
σ_A	5.2×10^{-3}	4.6×10^{-3}	5.0×10^{-3}	5.5×10^{-3}
$\sigma_{4,A}$	5.0×10^{-3}	4.5×10^{-3}	4.9×10^{-3}	5.3×10^{-3}
ρ_μ	0.64	0.58	0.61	0.64
σ_μ	4.0×10^{-2}	4.0×10^{-2}	4.3×10^{-2}	4.6×10^{-2}
$c_A(1)$	-0.013	-0.013	-0.012	-0.012
$d(1)$	0.31	0.27	0.29	0.31
ρ_p	0.89	0.88	0.90	0.91
μ_p	1.35	1.17	1.27	1.42
σ_p	0.67	0.57	0.65	0.75

Table 18 Estimates based on 10000 draws with 5000 burn-in.

1.6.7. Prior Distributions and Posterior Estimates

Parameter	Prior Dist	Para(1)	Para(2)
ρ_A	B	2.625	2.625
ρ_μ	B	2.625	2.625
ρ_{Π^*}	B	2.625	2.625
ρ_g	B	2.625	2.625
ρ_ψ	B	2.625	2.625
$c_A(s_t = 1)$	U	-	-
σ_A	U	-	-
$\sigma_{4,A}$	U	-	-
σ_μ	U	-	-
σ_{Π^*}	U	-	-
σ_g	U	-	-
σ_r	U	-	-
σ_ψ	U	-	-
ρ_p	U	0	1
μ_p	U	0	1
σ_p	U	-	-
Spp	N	3	1
ϕ_p	N	50	150
ϕ_w	N	1000	3000
γ_R	B	13.3125	4.4375
γ_Π	N	1.7	0.3
γ_y	G	1.44	0.12
h	B	12	12

Table 19 Downturn risk shocks model: NOTE: $c_A(1) = (1 - \phi_A)\mu_A(1)$. Prior distributions (B: Beta, G: Gamma, N: Normal, U: Uniform). Para(1) and Para(2) are two shape parameters for Beta distribution, shape and scale for Gamma distribution, mean and standard deviation for Normal distribution, lower bound and upper bound for Uniform distribution.

Parameter	Posterior Lower (10%)	Mean	Upper (90%)
ρ_A	0.47	0.53	0.59
ρ_μ	0.56	0.63	0.68
ρ_{Π^*}	0.98	0.99	0.99
ρ_g	0.98	0.99	0.99
ρ_ψ	0.27	0.35	0.43
$c_A(s_t = 1)$	-1.1×10^{-2}	-8.7×10^{-3}	-6.1×10^{-3}
$c_A(s_t = 2)$	1.7×10^{-3}	2.3×10^{-3}	2.9×10^{-3}
$d(s_t = 1)$	0.71	0.99	1.22
$d(s_t = 2)$	-0.34	-0.26	-0.20
σ_A	4.5×10^{-3}	5.1×10^{-3}	5.8×10^{-3}
$\sigma_{4,A}$	4.4×10^{-3}	5.0×10^{-3}	5.9×10^{-3}
σ_μ	5.4×10^{-2}	6.9×10^{-2}	8.8×10^{-2}
σ_{Π^*}	1.2×10^{-3}	1.4×10^{-3}	1.7×10^{-3}
σ_g	3.0×10^{-3}	3.2×10^{-3}	3.4×10^{-3}
σ_r	2.3×10^{-3}	2.5×10^{-3}	2.7×10^{-3}
σ_ψ	2.73	3.87	5.48
ρ_p	0.85	0.87	0.89
μ_p	0.84	1.34	1.70
σ_p	0.63	0.75	0.93
S''	2.47	3.26	4.31
ϕ_p	1098.7	1206.7	1329.8
ϕ_w	5550.3	7660.1	10651.8
γ_R	0.80	0.82	0.84
γ_Π	1.60	1.86	2.16
γ_y	0.70	0.82	0.96
h	0.78	0.80	0.83

Table 20 Downturn risk shocks model: NOTE: $c_A(1) = (1-\phi_A)\mu_A(1), c_A(2) = -\frac{(1-p)c_A(1)}{p}$.

1.6.8. Variance decompositions

Variance decomposition distributions constructed by taking every 10th draw of parameters after a burn-in of 50000 and simulating data series of length 2000. This results in 3000 samples.

Business cycle freq $(8-32$ qtrs)

	Downturn risk		Pure downturn risk	
	10%	90%	10%	90%
Out	24	37	22	34
Cons	52	64	21	31
Inv	5	13	9	20
Hours	19	29	25	42

Medium cycle freq $(32-200$ qtrs)

	Downturn risk		Pure downturn risk	
	10%	90%	10%	90%
Out	46	65	7	14
Cons	66	83	3	5
Inv	22	41	11	26
Hours	11	28	31	62

Business cycle freq $(8-32$ qtrs)

	Labor supply		Prod gr news	
	10%	90%	10%	90%
Out	1	2	2	5
Cons	0	1	6	11
Inv	1	1	1	3
Hours	2	5	4	7

Medium cycle freq $(32-200$ qtrs)

	Labor supply		Prod gr news	
	10%	90%	10%	90%
Out	1	3	6	11
Cons	0	1	5	9
Inv	3	6	7	13
Hours	10	18	1	3

1.6.9. Posterior predictive checks

Posterior predictive distributions constructed by taking every 10th draw of parameters after a burn-in of 50000 and simulating data series of length 208 (the sample size). This results in 1500 samples.

Standard deviation (×100)

Variable	Data	10%	Median	90%
Output gr	1.08	1.57	1.79	2.05
Consumption gr	0.53	0.92	1.10	1.32
Investment gr	3.38	3.97	4.44	5.00
Hours	4.94	6.27	8.41	11.67
Real wage gr	0.71	1.06	1.20	1.37
Inflation	0.59	0.95	1.30	1.87
Federal funds rate	0.87	0.94	1.30	1.87

Relative standard deviation to output growth

Variable	Data	10%	Median	90%
Consumption gr	0.49	0.53	0.62	0.70
Investment gr	3.14	2.27	2.50	2.70
Hours	4.59	3.53	4.70	6.36
Real wage gr	0.66	0.59	0.67	0.76
Inflation	0.55	0.52	0.72	1.07
Federal funds rate	0.81	0.52	0.72	1.06

Correlation with output growth

Variable	Data	10%	Median	90%
Consumption gr	0.64	0.47	0.61	0.72
Investment gr	0.94	0.87	0.90	0.93
Hours	0.15	−0.02	0.16	0.33
Real wage gr	0.14	0.27	0.41	0.54
Inflation	−0.23	−0.33	−0.12	0.10
Federal funds rate	−0.08	−0.08	0.14	0.33

1st order autocorrelation

Variable	Data	10%	Median	90%
Output gr	0.23	0.56	0.65	0.73
Consumption gr	0.42	0.78	0.84	0.89
Investment gr	0.20	0.47	0.56	0.64
Hours	0.97	0.95	0.97	0.98
Real wage gr	−0.04	0.46	0.58	0.68
Inflation	0.88	0.91	0.95	0.97
Federal funds rate	0.96	0.93	0.96	0.98

Because skewness of a data series is greatly impacted by small samples, for these calculations, I simulate paths of length 25000 for each parameter draw.

Skewness

Variable	Data	10%	Median	90%
Output gr	−0.64	−0.21	−0.14	−0.07
Consumption gr	−1.00	−0.59	−0.48	−0.25
Investment gr	−0.50	−0.05	−0.02	0.01
Labor	−0.73	−0.13	−0.04	0.05
Real wage gr	−0.02	−0.22	−0.15	−0.07
Inflation	1.24	−0.13	0.01	0.14
Federal funds rate	0.89	−0.13	0.00	0.13

CHAPTER 2 : A New Approach to Identifying the Real Effects of Uncertainty Shocks

2.1. Introduction

This paper proposes and estimates a multivariate stochastic volatility in vector autoregression model called the conditional autoregressive inverse Wishart-in-vector autoregression (CAIW-in-VAR) model. The model allows for a first-order effect of stochastic volatility, which gives a theoretically-consistent framework to estimate and evaluate the real effects of uncertainty shocks. We also propose a novel identification strategy of uncertainty shocks that does not rely on the identification of level shocks. Therefore, we can make statements about identified uncertainty shocks without needing to risk misspecification in the level innovations. We further provide algorithms to compute impulse response functions and variance decompositions. To illustrate our model, we present an empirical exercise evaluating the real effects of financial uncertainty shocks. Our main result is that uncertainty shocks originating in the financial sector, as opposed to those from the real sector, generate the most important first-order effects. In fact, an increase in financial uncertainty leads to a permanent decline in real activity.

Starting with work by Bloom (2009) and Fernandez-Villaverde, Guerron-Quintana, Rubio-Ramirez, and Uribe (2011), researchers have become interested in the first-order effects of uncertainty shocks. Much of the research in the vector autoregression framework, however, has used proxies for volatility, such as the VIX (Bloom (2009)) or the Economic Policy Uncertainty Index of Baker, Bloom, and Davis (2013). Identified innovations to this volatility proxy are then interpreted as volatility shocks and its real effects come from the impulse

responses of the other variables in the vector autoregression. In contrast to this framework, our model connects uncertainty to the stochastic volatility estimated in the variables. Therefore, we do not run into the issue of interpreting the uncertainty that the volatility proxy truly captures[1]. Moreover, the multivariate framework of our model allows us to identify multiple sources of uncertainty (i.e. monetary policy versus financial) in a theoretically-coherent fashion. We can then compare the responses of the macroeconomic variables to these uncertainty shocks and use forecast error variance decompositions to determine the relative importance of the real effects from these sources of uncertainty.

Our new identification strategy for the uncertainty shocks closely connects with our conditional autoregressive inverse Wishart volatility process. This volatility process, introduced into the financial econometrics literature by Golosnoy, Gribisch, and Liesenfeld (2012) and in macroeconomics by Karapanagiotidis (2012), models the time-varying volatility with the Wishart family of distributions (See Gourieroux, Jasiak, and Sufana (2009) and Fox and West (2013) for alternative autoregressive Wishart models.). The structure of the volatility dynamics makes it such that the process can be written in a linear vector autoregressive form with innovations that are martingale difference sequences. Upon writing the volatility process in vector autoregressive form, it is possible to use many of the identification schemes developed in the structural VAR framework without having to consider identification of the level innovations. We currently only discuss the recursive identification scheme and are considering extensions to other identification strategies.

We develop a Markov chain Monte Carlo algorithm relying on a Metropolis-within-Gibbs sampler to simulate from the posterior density. The algorithm builds off of the one developed in Philipov and Glickman (2006), Rinnergschwentner, Tappeiner, and Walde (2011), and Karapanagiotidis (2012). It is complicated by the fact that the stochastic volatility process also enters into the conditional mean equation, which the previous papers do not consider. In analyzing the estimated model, we also show how to construct impulse response functions

[1]For instance, it is unclear whether VIX measures macroeconomic uncertainty or just uncertainty in the financial sector.

to identified uncertainty shocks and forecast error variance decompositions.

As an empirical application, we investigate the relative importance of uncertainty originating in the financial sector and in the real sector. We estimate our model on industrial production growth in the manufacturing sector and the Chicago Fed National Financial Conditions Index. Identified innovations to the stochastic volatility in industrial production growth are real uncertainty shocks whereas innovations to the stochastic volatility in the Chicago Fed National Financial Conditions Index are financial uncertainty shocks. We find that an increase in financial uncertainty leads to a permanent decline in industrial production and a transitory worsening of financial conditions. Real uncertainty shocks, on the other hand, do not have significant effects on either industrial production or financial conditions. These facts show up in the variance decompositions as well. Our results indicate that around 20% of the forecast error variance of industrial production growth and 10% of the forecast error variance of financial conditions up to 10 years out come from uncertainty shocks. Among them, financial uncertainty shocks are dominant, accounting for around 90% of the first-order effects of uncertainty on industrial production growth and 60% for financial conditions.

Our work speaks to three main literatures. First, we contribute to the empirical literature investigating the real effects of uncertainty shocks. One strand of the literature, as mentioned already, uses identified innovations to volatility proxies in vector autoregressions to identify macroeconomic movements from uncertainty shocks. Another strand of literature, exemplified by Fernandez-Villaverde, Guerron-Quintana, Rubio-Ramirez, and Uribe (2011) and Fernandez-Villaverde, Guerron-Quintana, Kuester, and Rubio-Ramirez (2013), first estimates the stochastic volatility from a vector autoregression and then feeds the extracted process into a dynamic equilibrium model. Relative to this work, we propose a new model that can identify multiple sources of uncertainty shocks with no restrictions on the level innovations and can be estimated in one step. We hope this model will become a useful tool to guide researchers in determining the important sources of the real effects of uncertainty

in the economy.

The second literature with which we connect is the work on conditional heteroskedasticity-in-mean. French, Schwert, and Stambaugh (1987) first propose the GARCH-in-mean model in the financial economics context. Elder (2004), for inflation uncertainty, and Elder and Serletis (2010), for oil price uncertainty, use the model to investigate the volatility-in-mean effect in macroeconomic applications. Working with a stochastic volatility-in-mean model allows us to consider impulse responses to volatility shocks with no movements in the level innovations, which corresponds more closely to the notion of an uncertainty shock. Because GARCH-in-mean models do not have an independent source of variation driving volatility, they have a more cumbersome time producing these impulse responses. Koopman and Uspensky (2002) propose the univariate stochastic volatility-in-mean model. To answer many of our questions, a multivariate extension to the model is required, which we provide. Furthermore, we also propose a novel identification strategy for uncertainty shocks.

Finally, our modeling framework builds off of a line of research modeling vector autoregressions with time-varying volatility, beginning with Uhlig (1997). Other important contributions in this field include Cogley and Sargent (2005), Primiceri (2005), and Sims and Zha (2006). These papers are concerned with whether stochastic volatility, representing changes in the nature of shocks hitting the economy, or coefficient changes, representing shifts in the underlying relationships in the economy, are more responsible for the evolving nature of macroeconomic movements in the U.S. We also consider a vector autoregression with stochastic volatility, but we allow the stochastic volatility to have a conditional mean effect as well. Therefore, while our model builds off of this literature, the questions we aim to answer are quite different. In principle, we can allow for coefficient drift as well, but for clarity of presentation, we shut off that channel.

The plan of the paper is as follows. In section 2, we present the model and discuss identification of uncertainty shocks. In section 3, we introduce our Markov chain Monte Carlo sampler and our algorithms to compute impulse responses and forecast error variance de-

compositions. Section 4 contains our empirical application on the real effects of financial uncertainty shocks and section 5 concludes.

2.2. Model

In this section, we lay out our proposed Conditional Autoregressive Inverse Wishart-in-Vector Autoregression (CAIW-in-VAR) model. We begin by discussing our model specification. Then, we provide details on our volatility process and discuss the advantages of our modeling strategy. After that, we present our definition of an uncertainty shock. Finally, we introduce our novel identification strategy for uncertainty shocks.

2.2.1. Model specification

We consider the following vector autoregression with multivariate stochastic volatility,

$$Y_t = \mu + \Phi Y_{t-1} + Bf(\Sigma_t) + \epsilon_t, \quad \epsilon_t|\Sigma_t \sim N(0,\ \Sigma_t) \tag{2.1}$$

where Y_t and μ are $k \times 1$ vectors, Φ is a $k \times k$ matrix, B is a $k \times l$ matrix (or vector), and $f(\cdot)$ is a known function that maps a $k \times k$ matrix into an $l \times 1$ vector. The forecast error ϵ_t is conditionally multivariate normal with a $k \times k$ time-varying covariance matrix Σ_t.

The term $Bf(\Sigma_t)$ captures the phenomenon called the "real effect of uncertainty shock" in macroeconomics, which allows fluctuations in the volatility of the shocks to change the conditional mean of the process. In a structural model with optimizing agents, this effect would come from agents' optimal responses to changes in risk in the economy. In our paper, we present an econometric model that can allow for these effects.

Volatility process. We model the multivariate stochastic volatility with Wishart processes as in Philipov and Glickman (2006), Golosnoy, Gribisch, and Liesenfeld (2012), and Karapanagiotidis (2012),

$$\Sigma_{t+1}|\nu, S_t \sim IW(\nu, S_t^{-1}), \tag{2.2}$$

where $\nu > k+1$ is a scalar. The dynamics of the multivariate stochastic volatility are modelled by a $k \times k$ matrix S_t, which is defined with two additional parameter matrices C and A.

$$S_t = \frac{1}{(\nu - k - 1)}(C + A\Sigma_t A')^{-1} \tag{2.3}$$

C is a $k \times k$ positive definite matrix that governs the long-run mean of the multivariate volatility process. A is a $k \times k$ matrix that governs the dynamic properties of the volatility matrix process. This formulation ensures that the resulting scale matrix S_t is symmetric and positive definite.

Note that the process is formulated in a way that the conditional mean of the volatility matrix has the following simple form

$$E[\Sigma_{t+1}|\mathcal{F}_t] = C + A\Sigma_t A' \tag{2.4}$$

and

$$Cov(\Sigma_{ij,t+1},\ \Sigma_{lm,t+1}|\mathcal{F}_t) = \frac{2\Psi_{ij,t+1}\Psi_{lm,t+1} + (\nu - k + 1)(\Psi_{il,t+1}\Psi_{jm,t+1} + \Psi_{im,t+1}\Psi_{lj,t+1})}{(\nu - k)(\nu - k - 3)} \tag{2.5}$$

where $\mathcal{F}_t = \{\Sigma_t, \Sigma_{t-1}, ...\}$ and $\Psi_t = C + A\Sigma_t A'$. This delivers a convenient linear representation for the multivariate volatility process with innovations that are martingale difference sequences,

$$\sigma_t = \bar{C} + \bar{A}\sigma_{t-1} + \bar{v}_t, \quad E[\bar{v}_t] = 0 \quad \text{and} \quad E[\bar{v}_t \bar{v}_s'] = 0, \quad \forall s \neq t \tag{2.6}$$

where $\sigma_t = vech(\Sigma_t)$, $\overline{C} = vech(C)$ and

$$\bar{A} = L_n(A \otimes A)D_n$$

where $vec(x) = D_n vech(x)$ and $vech(x) = L_n vec(x)$.

The VAR form of the volatility process proves key to deriving unconditional moments and giving stationarity conditions, as discussed in Golosnoy, Gribisch, and Liesenfeld (2012). Moreover, for the purposes of this paper, we will discuss how this form greatly facilitates the identification of uncertainty shocks without needing to identify the level structural shocks.

Specifications for $f(\Sigma_t)$. The specification of the function $f(\Sigma_t)$ is left up to the researcher as long as $f(\Sigma_t)$ enters in the conditional mean equation linearly through B. We list a few specifications for $f(\Sigma_t)$: $diag(chol(\Sigma_t)), diag(\Sigma_t), \log(diag(\Sigma_t))$, and $vech(\Sigma_t)$. In our empirical application, we present results for $f(\Sigma_t) = diag(chol(\Sigma_t))$, noting that the qualitative results are robust to the other specifications.

2.2.2. Identification of Uncertainty Shocks

We define an *uncertainty shock to variable z* as a shock that increases the forecast error variance of variable z. An uncertainty shock to variable z makes it harder for economic agents to forecast variable z (agents become more uncertain about variable z). This definition of an uncertainty shock is related to Jurado, Ludvigson, and Ng (2013), where they define uncertainty as the conditional variance of the unforecastable component of an economic variable.

Recall the VAR representation of the CAIW process,

$$\sigma_t = \bar{C} + \bar{A}\sigma_{t-1} + \bar{v}_t, \quad E[\bar{v}_t] = 0 \quad \text{and} \quad E[\bar{v}_t \bar{v}_s'] = 0, \, \forall s \neq t, \quad E[\bar{v}_t \bar{v}_t'] = \Omega_t$$

where Ω_t is the conditional variance of σ_t given the information set at time $t-1$ and is a closed-form function of Σ_{t-1}. This VAR form in the stochastic volatilities crucially depends on our CAIW process of stochastic volatility. The VAR framework naturally allows us to use the identification strategies developed in the structural VAR literature to identify uncertainty shocks. Moreover, identifying the uncertainty shocks does not require identi-

fying any level structural shocks. Therefore, we can make statements about the impact of uncertainty shocks without needing to risk model misspecification in the level shocks.

Example: Recursive identification. As an illustrative example, we present a recursive identification scheme for a 2-variable CAIW-in-VAR model.

$$\begin{pmatrix} \Sigma_{11,t} \\ \Sigma_{12,t} \\ \Sigma_{22,t} \end{pmatrix} = \bar{C} + \bar{A} \begin{pmatrix} \Sigma_{11,t-1} \\ \Sigma_{12,t-1} \\ \Sigma_{22,t-1} \end{pmatrix} + \underbrace{\begin{pmatrix} L_{11,11,t} & 0 & 0 \\ L_{12,11,t} & L_{12,12,t} & 0 \\ L_{22,11,t} & L_{22,12,t} & L_{22,22,t} \end{pmatrix} \begin{pmatrix} v_{11,t} \\ v_{12,t} \\ v_{22,t} \end{pmatrix}}_{=\bar{v}_t} \tag{2.7}$$

v_t is the vector of identified uncertainty shocks from the identification scheme. The L_t matrix that multiplies the uncertainty shocks is a function of Σ_{t-1} and other model parameters. With this identification strategy, orthogonalized innovations to $\Sigma_{11,t}$ contemporaneously affect the entire variance covariance matrix. Innovations to $\Sigma_{12,t}$, or the covariance between the first and second variables also contemporaneously impact the variance of the second variable. Finally, innovations to $\Sigma_{22,t}$ do not have any contemporaneous effect on any other terms in the variance covariance matrix.

Relationship with real effect of uncertainty shock to a structural shock. Another popular definition of an uncertainty shock is an increase in the volatility of a structural shock in a business cycle model. While our model does not exactly nest the solution to a business cycle model with higher-order (beyond 3) perturbed terms, as we do not include quadratic or cubic terms in our model, our framework can nevertheless provide suggestive evidence as to the comovements the uncertainty shock to the structural shock should produce in the business cycle model [2].

The key element necessary to operationalize this identification strategy is to have a data counterpart to the business cycle model structural shock in the CAIW-in-VAR. For example,

[2]See Aruoba, Bocola, and Schorfheide (2013b) for a framework that can be used to evaluate the performance of dynamic equilibrium models in capturing nonlinearities in the data.

suppose we are interested in the real effects of a total factor productivity uncertainty shock. To evaluate these effects, augment the data of the endogenous variables of interest (for example, output, consumption, investment, and hours) with an observable total factor productivity series such as from Fernald (2012). The comovements from an uncertainty shock to total factor productivity produced by our estimated model could provide guidance as to the comovements one would like to see a total factor productivity uncertainty shock produce in the business cycle model. Note that such an identification strategy would require the identification of the level structural shock implicitly through the choice of observable series.

2.3. Bayesian Analysis of CAIW-in-VAR Models

The objective of this section is to describe the econometric techniques used to estimate and analyze the CAIW-in-VAR model presented in the previous section. We take a Bayesian perspective and provide a simulation-based method to obtain approximations to the posterior distribution of unknown parameters and associated functionals of interest, such as impulse response functions and variance decompositions. We start this section by specifying prior distributions on the unknown parameters. Then, we describe the posterior simulator for the model. Finally, we lay out algorithms to get the posterior distributions of impulse response functions and variance decompositions using draws from the posterior sampler.

2.3.1. Posterior inference

Prior specification. As we take a Bayesian perspective, the presented CAIW-in-VAR model is completed by specifying prior distributions on the unknown parameters. Parameters in the conditional mean of the model, μ, Φ and B, are assumed to follow independent multivariate normal distributions,

$$\mu \sim N\left(m_\mu, V_\mu\right), \quad vec(\Phi) \sim N\left(m_\Phi, V_\Phi\right), \quad vec(B) \sim N\left(m_B, V_B\right)$$

where $vec(\cdot)$ is the vectorize operator. The choice of this prior specification facilitates posterior computation due to its conjugacy.

There are three types of parameters in the volatility equation (A, C, and ν). The parameter A governs the dynamic properties of the volatility matrix process. Each element of A follows an independent normal distribution except the element in the far upper-left corner. The prior distribution for the $(1,1)$-th element in the A matrix is set to be a truncated normal distribution defined on the positive real line to ensure identification (see Golosnoy, Gribisch, and Liesenfeld (2012) for more details).

$$A(1,1) \sim TN(m_{A(1,1)}, V_{A(1,1)}, 0, \infty)$$
$$A(i,j) \sim N(m_{A(i,j)}, V_{A(i,j)}) \quad \forall\, (i,j) \neq (1,1)$$

The parameter C determines the long-run mean of the volatility process. We set the prior for it as following an inverse Wishart distribution with scale matrix Ψ and degrees of freedom parameter df. As the Wishart-type distribution is quite a popular prior in the Bayesian literature for a variance covariance matrix, we believe it to be a natural choice for C.

$$C \sim IW\,(df, \Psi)$$

Finally, ν, the degrees of freedom parameter, follows a uniform distribution with lower bound to ensure that the inverse Wishart process is well defined and the variances of its elements exist, $\nu > k + 3$.

$$\nu \sim U(k+3, M_\nu)$$

where M_ν is large but bounded real number.

Posterior simulator. We construct a Metropolis-within-Gibbs posterior simulator to draw from the posterior distribution of our parameters. The algorithm runs on the following cycles:

1. $p(\Sigma_t|\ others, Y)$ for $t = 1, ..., T-1$: multivariate stochastic volatilities
2. $p(\Sigma_T|\ others, Y)$: multivariate stochastic volatility at the last period
3. $p(\mu, B, \Phi|\ others, Y)$: parameters in the conditional mean equation
4. $p(\nu|\ others, Y)$: degrees of freedom parameter in the inverse Wishart process
5. $p(C|\ others, Y)$: long-run mean parameter in the inverse Wishart process
6. $p(A|\ others, Y)$: dynamics parameter in the inverse Wishart process

where we define $p(\theta|others, Y)$ as the conditional distribution of θ given $Y_{1:T}$ and all other parameters except θ. In the appendix, we provide details of the algorithm with full conditional posterior distributions.

Our algorithm builds upon the specifications of Philipov and Glickman (2006), Rinnergschwentner, Tappeiner, and Walde (2011), and Karapanagiotidis (2012). Relative to their frameworks, our model has a complication in that the stochastic volatility appears in the conditional mean equation as well and therefore their posterior samplers are not directly applicable to our framework. Our algorithm adopts the single-move state simulator of Jacquier, Polson, and Rossi (1994), which is widely used in the context of the univariate stochastic volatility model[3].

2.3.2. Impulse response function

Calculating the impulse response function to an identified uncertainty shock is complicated by the nonlinear nature of the system. Specifically, a critical nonlinearity is that Ω_t, the time

[3]See Cogley and Sargent (2005) and Clark (2011) for macroeconomic applications

t variance covariance matrix of $\bar{v}_t$, is a function of Σ_{t-1}. Therefore, as with all nonlinear impulse response functions, the response is not invariant to the initial value. Conditional upon a parameter draw, we use simulation methods to compute the impulse response function.

Continuing with our 2-variable CAIW-in-VAR recursive identification example in equation 2.7, we present how to calculate an impulse response to an uncertainty shock. An extension to a larger dimension is straightforward.

Algorithm 1. *(IRF of the uncertainty shock of length S in the 2-variable CAIW-in-VAR model recursive identification)*

1. *Choose initial value Σ_{-1} for IRF.*

2. *Consider 1 standard deviation increase of an element in v_0 (here we present a 1 standard deviation increase of the first element in the vector v_0).*

$$v_0 \sim (0, I)$$

$$v_0^1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \textit{versus} \quad v_0^0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

3. *Simulate two volatility paths conditional on initial shock.*

$$\{\Sigma_t(v_0^1)\}_{t=0,\ldots,S} \quad \textit{versus} \quad \{\Sigma_t(v_0^0)\}_{t=0,\ldots,S}$$

   Repeat M times.

4. *For each simulation, calculate the values $\{Y_t(v_0^1)\}_{t=0,\ldots,S}, \{Y_t(v_0^0)\}_{t=0,\ldots,S}$ implied by the volatility paths.*

5. Form impulse response:

$$E\left[\frac{\partial Y_S}{\partial \text{ uncertainty shock}}\right] \approx \frac{1}{M}\sum_{m=0}^{M} Y_{0:S}^{(m)}(v_0^1) - \frac{1}{M}\sum_{m=0}^{M} Y_{0:S}^{(m)}(v_0^0)$$

2.3.3. Variance decomposition

We discuss how to construct the variance decomposition. Computing the variance decomposition is important to estimate the quantitative importance of the identified uncertainty shocks to fluctuations. To compute the forecast error variance (Algorithm 3), we require the algorithm for the historical variance decomposition (Algorithm 2), which we develop as well. The historical variance decomposition determines the historical fluctuations that originated from the identified uncertainty shock. We continue with our 2-variable CAIW-in-VAR recursive identification example in equation 2.7.

Algorithm 2. *(Historical variance decomposition) Enter the algorithm with* $\{Y_t, \widehat{\Sigma}_t\}_{t=0}^T$ *and the parameter vector,* $(\widehat{\mu}, \widehat{\Phi}, \widehat{B}, \widehat{\overline{C}}, \widehat{\overline{A}})$.

1. Extract the historical values of $\widehat{\epsilon}_t$*:*

$$\widehat{\epsilon}_t = Y_t - (\widehat{\mu} + \widehat{\Phi} Y_{t-1} + \widehat{B} f(\widehat{\Sigma}_t)).$$

2. Extract the historical values of $\widehat{v}_t$*:*

$$\widehat{v}_t = \widehat{\Upsilon}_t^{-1}(\widehat{\sigma}_t - \widehat{\overline{C}} - \widehat{\overline{A}}\widehat{\sigma}_{t-1}) \quad \textit{for } t = 1, ..., T$$

where $\widehat{\sigma}_t = vech(\widehat{\Sigma}_t)$ *and* $\widehat{\Upsilon}_t = Chol(Cov(\sigma_t, \sigma_t | \widehat{\mathcal{F}}_t^i))$ *where* $\widehat{\mathcal{F}}_t^i = \{\widehat{\sigma}_{t-1}^i, \widehat{\sigma}_{t-2}^i, ...\widehat{\sigma}_0^i\}$.

3. For each uncertainty shock $i = 1, .., \frac{k(k-1)}{2}$*,*

(a) Define $v_t^{*,i}$ *as a vector with the same dimension as* v_t *by setting the i'th element* $v_t^{*,i} = \widehat{v}_t^i$*. All other elements of the vector* $v_t^{*,i}$ *are set equal to zero.*

(b) Starting from $\sigma_0^{*,i} = \widehat{\sigma}_0$, iterate the following equations forward:

$$\Upsilon_t^{*,i} = Chol\left(Cov(\sigma_t, \sigma_t | \mathcal{F}_t^{*,i})\right), \quad where \quad \mathcal{F}_t^{*,i} = \{\sigma_{t-1}^{*,i}, \sigma_{t-2}^{*,i}, ...\sigma_0^{*,i}\}$$
$$\sigma_t^{*,i} = \overline{C} + \overline{A}\sigma_{t-1}^{*,i} + \Upsilon_t^{*,i} v_t^{*,i}$$
$$Y_t^{*,i} = \widehat{\mu} + \widehat{\Phi} Y_{t-1}^{*,i} + \widehat{B} f(\Sigma_t^{*,i}), \quad where \quad \sigma_t^{*,i} = vech(\Sigma_t^{*,i})$$

and set $\Sigma_t^{*,i} = \Sigma_t^{*,i}$ and $Y_t^{*,i} = Y_t^{*,i}$ for $t = 1, ..., T$ and $i = 1, ..., \frac{k(k-1)}{2}$.

To compute the forecast error decomposition of the uncertainty shocks, we first generate multiple simulated paths of forecasts from the estimated model ($Y_t^{(m)}$ for simulation m). Then, we apply the historical variance decomposition (Algorithm 2) to decompose the variation of each simulated $Y_t^{(m)}$ from the identified uncertainty shocks. Finally, we compute the share of the forecast error variance of Y_t from the identified uncertainty shocks. We denote $VC_s^{j,i,l}$ as the proportion of the s-step-ahead forecast error variance of the j-th variable due to the i-th uncertainty shock at the l-th parameter draws from the posterior simulator.

Algorithm 3. *(S-step ahead forecast error variance decomposition)*

1. *Repeat the following recursion for $m = 1, ..., M$ using*

$$Y_T, \widehat{\Sigma}_T^{(l)} \quad and \quad (\widehat{\mu}^{(l)}, \widehat{\Phi}^{(l)}, \widehat{B}^{(l)}, \widehat{\overline{C}}^{(l)}, \widehat{\overline{A}}^{(l)}),$$

where l stands for l-th posterior draws.

   (a) *(Generate forecasts) Starting from $Y_T^{(m)} = Y_T$ and $\Sigma_T^{(m)} = \widehat{\Sigma}_T^{(l)}$, iterate the CAIW-in-VAR model forward to obtain $Y_{T+1:T+S}^{(m)}$ and $\Sigma_{T+1:T+S}^{(m)}$.*

   (b) *(Shock decompositions) Apply the historical uncertainty shock decomposition algorithm with $Y_{T:T+S}^{(m)}$, $\Sigma_{T:T+S}^{(m)}$, and $(\widehat{\mu}^{(l)}, \widehat{\Phi}^{(l)}, \widehat{B}^{(l)}, \widehat{\overline{C}}^{(l)}, \widehat{\overline{A}}^{(l)})$ to obtain $\Sigma_t^{*,i,(m)}$ and $Y_t^{*,i,(m)}$ for $t = T+1, ..., T+S$ and $i = 1, ..., \frac{k(k-1)}{2}$.*

(c) Store $\Sigma^{(m)}_{T+1:T+S}$, $Y^{(m)}_{T+1:T+S}$, $\Sigma^{*,i,(m)}_{T+1:T+S}$, *and* $Y^{*,i,(m)}_{T+1:T+S}$ *for* $i = 1, ..., \frac{k(k-1)}{2}$.

2. Compute the forecast error variance share:

$$VC^{j,i,l}_s = \frac{\widehat{Var}(Y^{*,i}_{j,T+s})}{\widehat{Var}(Y_{j,T+s})}$$

for $i = 1, .., \frac{k(k-1)}{2}, j = 1, ..., k$, *and,* $s = 1, ..., S$. $\widehat{Var}(X)$ *is the Monte Carlo approximation to the variance of* X *using the simulated forecasts for* $m = 1, ..., M$.

We are also interested in quantifying the importance of the identified uncertainty shocks relative to the total forecast error variance coming only from uncertainty shocks. To this end, we also compute the total s-step-ahead forecast error variance of the variable j only due to uncertainty shocks as $\sum_{i=1}^{k(k-1)/2} VC^{j,i,l}_s$. Then, the share of the i-th uncertainty shocks among the forecast error variance of the j-th variable due to all uncertainty shocks is

$$VC^{j,i,l}_{unc,s} = \frac{VC^{j,i,l}_s}{\sum_{i=1}^{k(k-1)/2} VC^{j,i,l}_s}.$$

2.4. Empirical Application: Real effects of financial volatility

Motivation. Since the Great Recession, a renewed interest has emerged on the role of the financial sector in macroeconomic fluctuations. Our empirical application evaluates the real effects of uncertainty originating in the financial sector. Heightened uncertainty for future financial conditions could interfere with banks' willingness to lend, thereby disrupting efficient capital flows. This could be a potential channel through which financial uncertainty would affect the real economy.

Several strands of evidence point to the importance of the first-order effects of financial uncertainty. First, financial sector shocks seem to be central in explaining macroeconomic

movements. Jermann and Quadrini (2012) and Christiano, Motto, and Rostagno (2014) find that shocks originating in the financial sector have significant macroeconomic effects. Second, there appears to be evidence of a large amount of uncertainty originating in the financial sector. Fernandez-Villaverde and Rubio-Ramirez (2007) and Justiniano and Primiceri (2008) observe that volatility from shocks to the capital accumulation equation, which may be interpreted as financial sector shocks (e.g. Justiniano, Primiceri, and Tambalotti (2011b)), are the main drivers of the time-varying volatility of macroeconomic fluctuations on U.S. data. Finally, the financial channel seems to be important in the transmission of uncertainty shocks. Using proxies for uncertainty and financial conditions, Caldara, Fuentes-Albero, Gilchrist, and Zakrajsek (2013) find that uncertainty shocks can lead to a significant recession only if they are transmitted through the financial channel.

What macroeconomic movements follow financial uncertainty shocks? How important are uncertainty shocks for macroeconomic movements relative to level shocks? Among uncertainty shocks, how important is financial uncertainty relative to real uncertainty? Using our new framework, we provide answers to these questions.

Data. We use monthly industrial production growth in the manufacturing sector and the Chicago Fed National Financial Conditions Index (NFCI) from $1973M1 - 2012M12$. We obtained the industrial production growth and the NFCI data from the Federal Reserve Bank of St. Louis FRED. The NFCI is an index of financial conditions constructed from 105 measures of financial activity. We view the NFCI as a measure of the state of the financial sector in the U.S. economy. Alessandri and Mumtaz (2014b) also use this index, noting that it provides a better forecasting power of future industrial production growth than even the Excess Bond Premium of Gilchrist and Zakrajsek (2012).

Figure 21 displays monthly manufacturing industrial production growth (top) and the Chicago Fed National Financial Conditions Index (bottom). IP growth declines drastically during recession periods, especially in the mid 1970s and early 1980s, as well as the

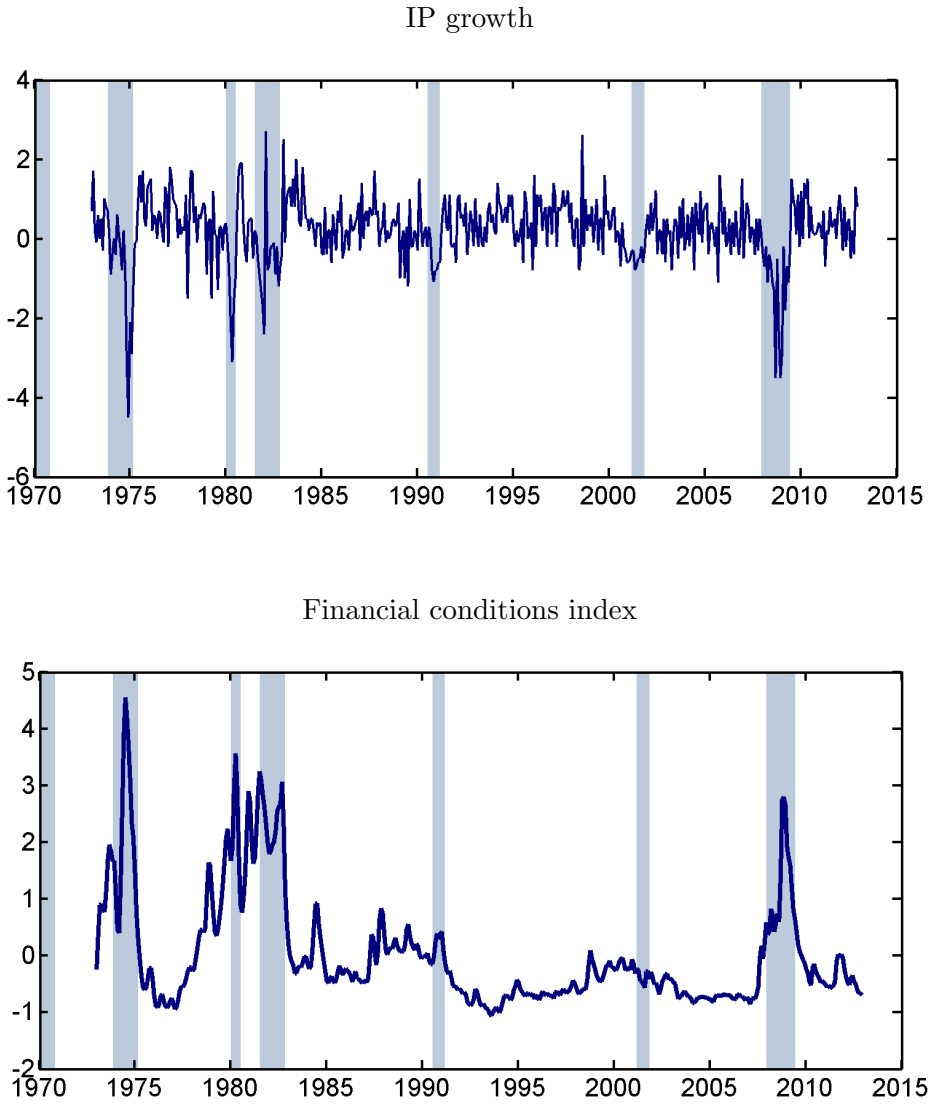


Figure 21 Monthly manufacturing industrial production growth (top) and Chicago Fed National Financial Conditions Index (bottom) $1973M1 - 2012M12$. Blue bars indicate NBER recession dates.

Great Recession. During the two recessions in the early 1990s and early 2000s, the IP growth declines were much milder. Note that financial distress seems to jump during the deep recessions of the 1970s, 1980s, and the Great Recession, but not so much during the mild recessions. Both series exhibit the well-known Great Moderation across the mid 1980s into the mid 2000s.

Model and prior specification. We use a 2-variable specification of our CAIW-in-VAR model. Our preferred function linking the volatility to the mean portion of the model is $f(\Sigma_t) = diag(chol(\Sigma_t))$, although our qualitative results are robust to a variety of linking functions[4]. We choose a VAR(4) for the lag length of the level variables, noting that this lag length essentially removes all autocorrelation from the fitted residuals in a model without stochastic volatility. For the CAIW process, we choose a lag length 1. In estimating the model, we take $100,000$ Markov chain Monte Carlo draws from our Gibbs sampler. We estimate all parameters in the model aside from ν, which we fix at 50[5]. Prior specifications and posterior intervals can be found in the appendix. Estimates of the smoothed stochastic volatility process are also in the appendix.

Real effect of uncertainty: IRF. We consider the real effects of real activity and financial conditions uncertainty shocks. To identify the uncertainty shocks, we use a recursive identification of the innovations in the vector autoregressive form of the volatility process with real activity uncertainty ordered first and financial uncertainty ordered last. The implication of this identification scheme is that a real activity uncertainty innovation contemporaneously impacts uncertainty in all aspects of the economy whereas a financial uncertainty innovation only contemporaneously impacts financial uncertainty. Note that these identifying restrictions do not have any implications for the level structural shocks. We initialize all impulse response functions at the long-run level of volatility. Although we only present figures for the impulse responses up to 40 months out, our results remain unchanged up to 120 months.

Consider the effects on industrial production of a real activity uncertainty shock and financial uncertainty shock (1 standard deviation increase) as shown in figure 22. The real activity uncertainty shock has no effect on industrial production. The median effect is small and the posterior bands indicate little significance to the results. On the other hand, a financial uncertainty shock has a highly significant permanent negative effect on industrial

[4]We also tried $f(\Sigma_t) = \log(diag(\Sigma_t)), diag(\Sigma_t)$, and $vech(\Sigma_t)$.
[5]The estimation currently has a hard time identifying ν.

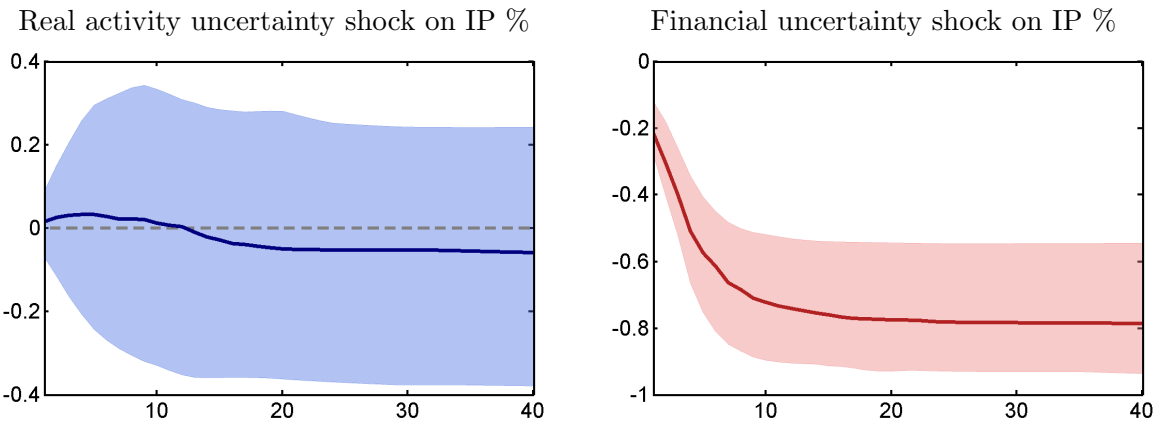


Figure 22 Effect of 1 standard deviation positive real activity uncertainty shock (left) and financial uncertainty shock (right) on IP level. Dark lines give posterior median results. Bands indicate 80% posterior intervals. Time period is in months. Results remain unchanged up to 120 months out.

production. The effect is -0.2% on impact and the long-run effect is -0.8%. There are two main takeaways from these results. First, upon controlling for financial uncertainty, real activity uncertainty has little empirical importance. Therefore, these findings suggest that a fruitful area for structural macroeconomic research is to consider models where uncertainty on structural shocks in the financial markets have large macroeconomic effects. Second, financial uncertainty shocks have a permanent impact on real activity. The previous literature on the real effects of uncertainty shocks has largely focused on the transitory real effects of uncertainty shocks. In contrast, our results suggest that increased financial uncertainty can have long-horizon real effects as well.

Figure 23 shows the response of financial conditions to a 1 standard deviation increase in real activity uncertainty and financial uncertainty. Similar to the earlier results, a real activity uncertainty shock does not have a significant impact on financial conditions. As the right panel shows, however, the financial uncertainty shock worsens financial conditions in a transitory, hump-shaped, fashion that lasts for around 15 quarters.

One potential explanation to justify these comovements is that increased financial uncertainty worsens financial conditions, which triggers capital misallocation due to the reduced

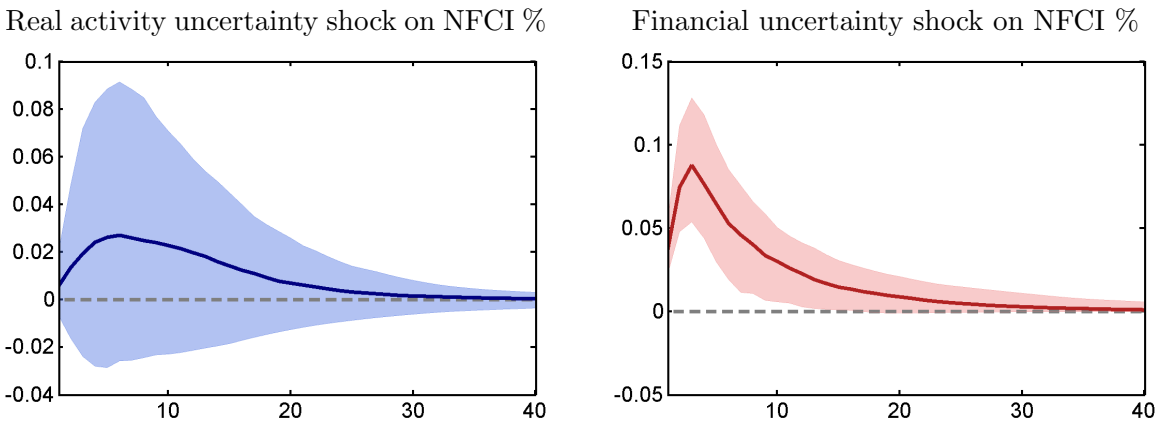


Figure 23 Effect of 1 standard deviation positive real activity uncertainty shock (left) and financial uncertainty shock (right) on NFCI. Dark lines give posterior median results. Bands indicate 80% posterior intervals. Time period is in months. Results remain unchanged up to 120 months out.

ability of the financial system to efficiently intermediate capital. Capital misallocation would show up in the aggregate as a decline productivity, which could lead to permanent real effects.

Forecast error variance decomposition. How important are the real effects of uncertainty shocks to macroeconomic fluctuations? To answer this question, we perform forecast error variance decompositions by month. Uncertainty shocks explain in total around 20% of industrial production growth forecast variance and 10% of financial conditions forecast variance. These values do not change much by month up to 120 months out. Figure 24 shows the percentage among the total fluctuations in industrial production growth accounted for by uncertainty shocks coming from real activity uncertainty and financial uncertainty. Real activity and financial uncertainty account for about the same amount of fluctuations upon impact. By 1 year out, however, real activity uncertainty essentially accounts for no industrial production growth movements, whereas financial uncertainty dominates the total contribution of uncertainty shocks. Similarly for financial stress, the real activity and financial uncertainty shocks account for around the same amount of fluctuations upon impact. Within half-a-year, however, financial uncertainty again dominates, accounting for around

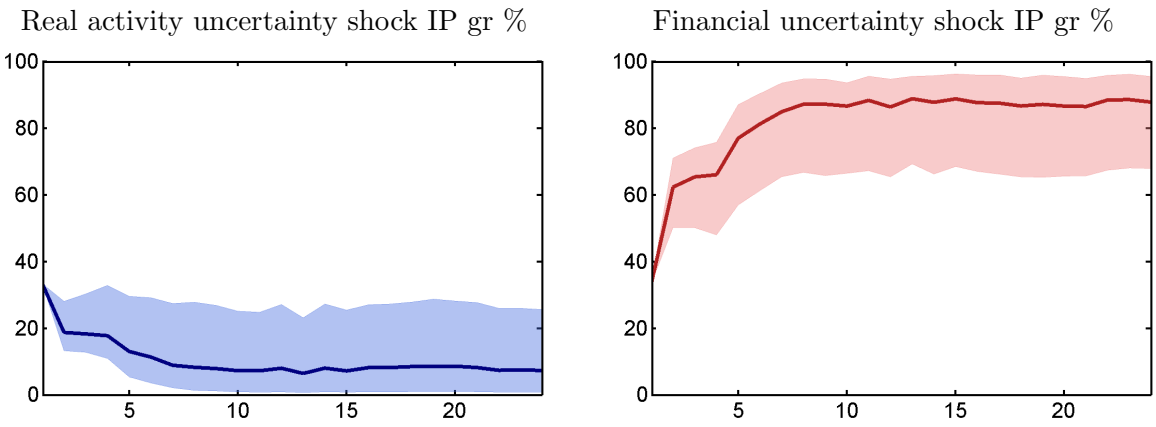


Figure 24 Percentage of total forecast error variance of industrial production growth from uncertainty shocks explained by real activity uncertainty (left) and financial uncertainty (right) by month. Dark lines give posterior median values while bands indicate 80% posterior intervals. Percentages do not add up to 100% because of effects of correlation shock between real activity and financial stochastic volatility. Time period is in months. Results remain unchanged up to 120 months out.

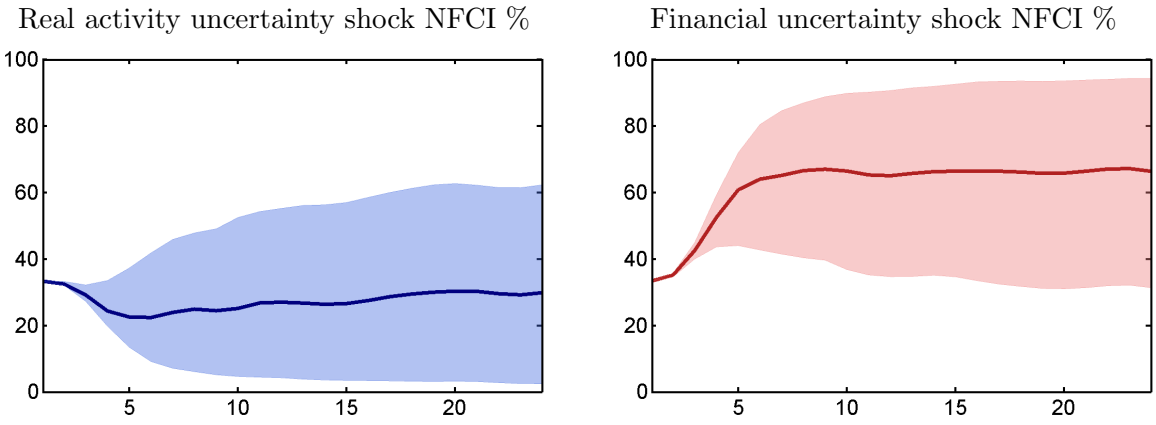


Figure 25 Percentage of total forecast error variance of NFCI from uncertainty shocks explained by real activity uncertainty (left) and financial uncertainty (right) by month. Dark lines give posterior median values while bands indicate 80% posterior intervals. Percentages do not add up to 100% because of effects of correlation shock between real activity and financial stochastic volatility. Time period is in months. Results remain unchanged up to 120 months out.

60% of fluctuations whereas real activity uncertainty only accounts for around 30%. These values stay stable up to 120 months out.

The variance decomposition results mirror the impulse response functions. Because a financial uncertainty shock appears to have a permanent effect on industrial production whereas a real activity uncertainty shock does not, financial uncertainty dominates the total contribution from uncertainty shocks in the long run. For the NFCI, the stronger response to a financial uncertainty shock relative to a real activity uncertainty shock leads to financial uncertainty again becoming the predominant driving force of uncertainty fluctuations.

2.5. Conclusion and Future direction

We have advanced the conditional autoregressive inverse Wishart-in-vector autoregression model to evaluate the real effects of uncertainty shocks. We discuss a novel identification strategy for uncertainty shocks that does not rely on the identification of level shocks. The paper also presents algorithms to construct impulse responses to uncertainty shocks and forecast error variance decompositions. In an empirical application, we evaluate the real effects of financial uncertainty shocks. Our results show that financial uncertainty shocks are an important driver of macroeconomic fluctuations, accounting for around 20% of industrial production movements. An increase in financial uncertainty leads to a permanent decline in industrial production and a transitory worsening of financial conditions. Uncertainty emerging in the real sector has comparatively small effects. These results suggest looking more closely at structural models with financial sector uncertainty.

In future work, we plan on investigating the viability of other identification strategies. These include long-run restrictions and sign restrictions. Long-run restrictions could represent a long-run shift of the volatility in the economy, such as the Great Moderation. Sign restrictions are quite natural in this framework, as oftentimes we are just interested in the effects of an increase in the uncertainty of a variable without wanting to take a stand on the recursive ordering of uncertainty innovations.

Finally, we are also interested in the forecasting performance of these models. The significant first-order effects of volatility changes leads us to believe that our class of models could outperform a standard vector autoregression with time-varying volatility in a forecasting horserace, especially around times of high volatility such as recessions.

2.6. Appendix

2.6.1. Details of the posterior sampler

The algorithm runs on the following cycles:

1. $p(\Sigma_t|\ others)$ for $t = 2, ..., T$: : SVs (require a modification)
2. $p(\Sigma_T|\ others)$: SV at the last period (require a modification)
3. $p(\mu, B, \Phi|\ others)$: parameter in the conditional mean equation (NEW)
4. $p(\nu|\ others)$: parameter in Wishart process (same as in original algorithm)
5. $p(C|\ others)$: parameter in Wishart process (same as in original algorithm)
6. $p(A|\ others)$: parameter in Wishart process (same as in original algorithm)

where $p(\theta|others)$ means the conditional distribution of θ given $Y_{1:T}$ and all other parameters except θ. We denote the previous draw as θ^{old}.

Note that the joint posterior distribution is

$$p(\mu, B, \Phi, \nu, C, A, \Sigma_{1:T}|Y_{t:T}) \propto p(Y_{t:T}|\mu, B, \Phi, \Sigma_{2:T})p(\Sigma_{2:T}|\nu, C, A, \Sigma_1)p(\mu, B, \Phi, \nu, C, A, \Sigma_1)$$

where the likelihood function can be decomposed as

$$p(Y_{2:T}|\mu, B, \Phi, \Sigma_{2:T}) = \prod_{t=2}^{T} p(Y_t|Y_{t-1}, \mu, B, \Phi, \Sigma_t).$$

To be able to construct an efficient MCMC sampling algorithm, we break the joint posterior into multiple blocks. For example, the conditional posterior distribution for the multivariate

stochastic volatility is decomposed into the following pieces[6],

$$p(\Sigma_{2:T}|\nu, C, A, \Sigma_1, Y_{2:T}) = \left(\prod_{t=2}^{T} p(\Sigma_t|\Sigma_{t-1}, \nu, C, A, Y_t) \right)$$

and iteratively sample Σ_t from $t = 2$ to $t = T$.

Step 1: Σ_t **for** $t = 2, ..., T$ The conditional posterior is

$$\begin{aligned} p(\Sigma_t|others) \propto & |\Sigma_t|^{-(\nu+k+1)/2} \exp\left(-\frac{1}{2} tr(S_{t-1}^{-1}\Sigma_t^{-1}) \right) \\ & \times |\Sigma_t|^{-1/2} \exp\left(-\frac{1}{2} tr\left(\epsilon_t' \Sigma_t^{-1} \epsilon_t\right) \right) \\ & \times |S_t|^{-\nu/2} \exp\left(-\frac{1}{2} tr\left(S_t^{-1}\Sigma_{t+1}^{-1}\right) \right) \end{aligned}$$

where we write

$$\epsilon_t = \underbrace{Y_t - \mu - \Phi Y_{t-1}}_{=e_t} - Bf(\Sigma_t) = e_t - Bf(\Sigma_t).$$

Then, we re-write the conditional posterior as

$$\begin{aligned} p(\Sigma_t|others) \propto & IW(\Sigma_t|\nu, \widetilde{S}_{t-1}^{-1}) \quad : \text{proposal density} \\ & \times |\Sigma_t|^{-1/2} |S_t|^{-\nu/2} \exp\left(-\frac{1}{2} tr\left(S_t^{-1}\Sigma_{t+1}^{-1}\right) \right) \quad : \text{MH correction 1} \\ & \times \exp\left(-\frac{1}{2} tr\left(\left[-2e_t g(\Sigma_t)' + g(\Sigma_t) g(\Sigma_t)' \right] \Sigma_t^{-1} \right) \right) \quad : \text{MH correction 2} \end{aligned}$$

where

$$g(\Sigma_t) = Bf(\Sigma_t) \quad \text{and} \quad \widetilde{S}_{t-1} = \left(S_{t-1}^{-1} + e_t e_t' \right)^{-1}.$$

We draw Σ_t based on the independent Metropolis-Hastings algorithm with the inverse Wishart distribution as a proposal distribution, $\Sigma_t^* \sim IW\left(\nu, \widetilde{S}_{t-1}^{-1}\right)$. The acceptance ratio

[6] Current version of the sampler fixes the initial stochastic volatility at the unconditional mean

$$E[\Sigma_1] = E[\Sigma_t] = (I - \bar{A})^{-1}\bar{C}.$$

More plausible implementation is to impose the inverse Wishart prior on Σ_1 centered around $E[\Sigma_t]$. Our current implementation can be viewed as a special case of this inverse Wishart prior.

is then,

$$r_{\Sigma_t}=\frac{|\Sigma_t^*|^{-1/2}|S_t^*|^{-\nu/2}\exp\left(-\frac{1}{2}tr\left((S_t^*)\Sigma_{t+1}^{-1}\right)\right)\exp\left(-\frac{1}{2}tr\left(\left[-2e_tg(\Sigma_t^*)'+g(\Sigma_t^*)g(\Sigma_t^*)'\right](\Sigma_t^*)^{-1}\right)\right)}{|\Sigma_t^{old}|^{-1/2}|S_t^{old}|^{-\nu/2}\exp\left(-\frac{1}{2}tr\left((S_t^{old})\Sigma_{t+1}^{-1}\right)\right)\exp\left(-\frac{1}{2}tr\left(\left[-2e_tg(\Sigma_t^{old})'+g(\Sigma_t^{old})g(\Sigma_t^{old})'\right](\Sigma_t^{old})^{-1}\right)\right)}$$

and we set $\Sigma_t^{new}=\Sigma_t^*$ with probability $max(r_{\Sigma_t},1)$, $\Sigma_t^{new}=\Sigma_t^{old}$ otherwise.

Step 2: Σ_T The full conditional posterior is

$$\begin{aligned}
p(\Sigma_T|others) &\propto |\Sigma_T|^{-(\nu+k+1)/2}\exp\left(-\frac{1}{2}tr\left(S_{T-1}^{-1}\Sigma_T^{-1}\right)\right)\times|\Sigma_T|^{-1/2}\exp\left(-\frac{1}{2}\epsilon_T'\Sigma_T^{-1}\epsilon_T\right)\\
&\propto |\Sigma_T|^{-(\nu+k+1)/2}\exp\left(-\frac{1}{2}tr\left(S_{T-1}^{-1}\Sigma_T^{-1}\right)\right)\times|\Sigma_T|^{-1/2}\exp\left(-\frac{1}{2}e_T'\Sigma_T^{-1}e_T\right)\\
&\quad\times\exp\left(-\frac{1}{2}\left(\epsilon_T'\Sigma_T^{-1}\epsilon_T-e_T'\Sigma_T^{-1}e_T\right)\right)\\
&\propto IW\left(\Sigma_T^{-1}\middle|\,\nu+1,\left(S_{T-1}^{-1}+e_Te_T'\right)\right)\quad:\text{ proposal density}\\
&\quad\times\exp\left(-\frac{1}{2}\left(\epsilon_T'\Sigma_T^{-1}\epsilon_T-e_T'\Sigma_T^{-1}e_T\right)\right)\quad:\text{ MH correction}
\end{aligned}$$

where we define e_T in the same as before. We draw Σ_T based on the independent Metropolis-Hastings algorithm with the inverse Wishart distribution as a proposal distribution, $\Sigma_T^*\sim IW\left(\nu+1,\left(S_{T-1}^{-1}+e_Te_T'\right)\right)$ and therefore the acceptance ratio is

$$r_{\Sigma_T}=\frac{\exp\left(-\frac{1}{2}\left((\epsilon_T^*)'(\Sigma_T^*)^{-1}\epsilon_T^*-e_T'(\Sigma_T^*)^{-1}e_T\right)\right)}{\exp\left(-\frac{1}{2}\left((\epsilon_T^{old})'(\Sigma_T^{old})^{-1}\epsilon_T^{old}-e_T'(\Sigma_T^{old})^{-1}e_T\right)\right)}$$

where

$$\epsilon_T^*=\underbrace{Y_T-\mu-\Phi Y_{T-1}}_{=e_T}-Bf(\Sigma_T^*)=e_T-Bf(\Sigma_T^*).$$

We set $\Sigma_T^{new}=\Sigma_T^*$ with probability $max(r_{\Sigma_T},1)$, $\Sigma_T^{new}=\Sigma_T^{old}$ otherwise.

Step 3: (μ,B,Φ) First we transform our model into the following multiple regression form,

$$\widetilde{Y}_t=\widetilde{B}\widetilde{X}_t+\Sigma_t^{1/2}\epsilon_t,\quad\epsilon_t\sim N(0,I)$$

where p is the number of lags in VAR and

$$\begin{aligned}\widetilde{Y}_t &= Y_t' \\ \widetilde{X}_t &= [1, Y_{t-1}', ..., Y_{t-p}', f(\Sigma_t)']' \\ \widetilde{B} &= [\mu, \Phi_1, ..., \Phi_p, B].\end{aligned}$$

Then, we can re-write the equation as

$$\Sigma_t^{-1/2}\widetilde{Y}_t = \left(\widetilde{X}_t' \otimes \Sigma_t^{-1/2}\right) vec\left(\widetilde{B}\right) + \epsilon_t,$$

which is a standard multiple regression with homoscedastic errors. The conditional posterior distribution of (μ, B, Φ) is a multivariate normal distribution under the conjugate prior assumption.

Step 4: ν The conditional posterior distribution of ν is

$$p(\nu|others) \propto \left(\prod_{t=2}^{T} \frac{|S_{t-1}^{-1}|^{\nu/2}}{2^{\nu k/2}\Gamma_k(\nu/2)} |\Sigma_t|^{-(\nu+k+1)/2} exp\left(-1/2 tr(S_{t-1}^{-1}\Sigma_t^{-1})\right)\right) \mathbf{1}_{(k+1,M_\nu)}(\nu)$$

where $S_{t-1}^{-1} = (v-k-1)(C + A\Sigma_{t-1}A')$, $\Gamma_k(\cdot)$ is the multivariate gamma function, and $\mathbf{1}_{(k+1,M_\nu)}(\nu)$ is a indicator function takes value 1 if $\nu \in (k+1, M_\nu)$ and 0 otherwise. To draw ν from this conditional posterior distribution, we employ the random-walk Metropolis-Hastings algorithm with a proposal

$$\nu^* = \nu^{old} + e_\nu, \quad e_\nu \sim N(0, \sigma_\nu^2)$$

where the scale of the proposal distribution σ_ν^2 is adaptively chosen so that the resulting acceptance rate is about 30% (Atchadé and Rosenthal, 2005).

Step 5: C The conditional posterior distribution of C is

$$p(C|others) \propto \left(\prod_{t=2}^{T} |S_{t-1}^{-1}|^{\nu/2} |\Sigma_t|^{-(\nu+k+1)/2} exp\left(-1/2 tr(S_{t-1}^{-1}\Sigma_t^{-1})\right) \right) \times p_{IW}(C|df, \Psi)$$

where $S_{t-1}^{-1} = (v-k-1)(C + A\Sigma_{t-1}A')$ and p_{IW} is a density function of the inverse Wishart distribution. In this step, we reparametrize C in the following fashion,

$$C = \begin{pmatrix} d_{11} & 0 & \dots & 0 \\ c_{21} & d_{22} & \dots & 0 \\ \vdots & & \ddots & 0 \\ c_{k1} & c_{k2} & c_{k3} & \dots d_{kk} \end{pmatrix} \begin{pmatrix} d_{11} & 0 & \dots & 0 \\ c_{21} & d_{22} & \dots & 0 \\ \vdots & & \ddots & 0 \\ c_{k1} & c_{k2} & c_{k3} & \dots d_{kk} \end{pmatrix}' .$$

This transformation ensures the positive definiteness of C. To draw C from this conditional posterior distribution, we employ the random-walk Metropolis-Hastings algorithm with a proposal

$$c_{ij}^{*} = c_{ij}^{old} + e_{c(i,j)}, \quad e_{c(i,j)} \sim N(0, \sigma_{c(i,j)}^2)$$
$$\log(d_{ii}^{*}) = \log(d_{ii}^{old}) + e_{d(i,i)}, \quad e_{d(i,i)} \sim N(0, \sigma_{d(i,i)}^2)$$

for $(i,j) = \{i = 1, .., k; j = 1, ..., k, i \geq j\}$. The scale of the proposal distribution $\sigma_{A(i,j)}^2$ is adaptively chosen so that the resulting acceptance rate is about 30% (Atchadé and Rosenthal, 2005). Note that to compute the acceptance ratio, we need a Jacobian term due to reparametrization,

$$|J| = \underbrace{2^k \prod_{i=1}^{k} d_{ii}^{k+1-i}}_{cholsky\ decomp.} \times \underbrace{\prod_{i=1}^{k} d_{ii}}_{log\ trans.} .$$

Step 6: A The conditional posterior distribution of A is

$$p(A|others) \propto \left(\prod_{t=2}^{T} |S_{t-1}^{-1}|^{\nu/2} |\Sigma_t|^{-(\nu+k+1)/2} exp\left(-1/2 tr(S_{t-1}^{-1}\Sigma_t^{-1})\right) \right) \times p_{TN}(A_{11}|m_{A(1,1)}, V_{A(1,1}, 0, \infty) \prod_{(i,j)\neq(1,1)} p_N(A_{ij}, m_{A(i,j)}, V_{A(i,j)})$$

where $S_{t-1}^{-1} = (v-k-1)(C + A\Sigma_{t-1}A')$, p_{TN} is a density function of the truncated normal distribution, and p_N is a density function of the normal distribution. To draw A from this conditional posterior distribution, we employ the random-walk Metropolis-Hastings algorithm with a proposal

$$A^* = A^{old} + \Omega$$

where Ω is a symmetric matrix and each elements are drawn from the normal distribution,

$$\Omega = \begin{pmatrix} w_{11} & \dots & w_{k1} \\ \vdots & \ddots & \vdots \\ w_{k1} & \dots & w_{kk} \end{pmatrix}, \qquad w_{ij} \sim N(0, \sigma^2_{A(i,j)})$$

where the scale of the proposal distribution $\sigma^2_{A(i,j)}$ is adaptively chosen so that the resulting acceptance rate is about 30% (Atchadé and Rosenthal, 2005).

2.6.2. Prior specification

In this section, we present the prior distributions used for the application section. Our benchmark model is the CAIW(1)-in-VAR(4) model with $diag(chol(\Sigma_t))$ as a linking function. Parameters in the conditional mean of the model, μ, Φ and B, are assumed to follow independent multivariate normal distributions,

$$\mu \sim N\left(0_2,\ 10^2 \cdot I_2\right), \quad vec(\Phi) \sim N\left(0_{16},\ 10^2 \cdot I_{16}\right), \quad vec(B) \sim N\left(0_4,\ 10^2 \cdot I_4\right)$$

where $vec(\cdot)$ is the vectorize operator, $0_\#$ is a $\# \times 1$ vector of zeros, and $I_\#$ is a $\# \times \#$ identity matrix.

There are three types of parameters in the volatility equation (A, C, and ν). The parameter A governs the dynamic properties of the volatility matrix process. Each element of A follows an independent normal distribution except the element in the far upper-left corner. The prior distribution for the $(1,1)$-th element in the A matrix is set to be a truncated normal distribution defined on the positive real line to ensure identification.

$$A(1,1) \sim TN(0.9,\ 0.1^2, 0, \infty), \quad A(2,2) \sim N(0.9,\ 0.1^2)$$
$$A(1,2) \sim N(0, 10^2), \quad A(2,1) \sim N(0, 10^2).$$

The parameter C determines the long-run mean of the volatility process. We set the prior for it as following an inverse Wishart distribution with scale matrix Ψ and degrees of freedom parameter df. As the Wishart-type distribution is quite a popular prior in the Bayesian literature for a variance covariance matrix, we believe it to be a natural choice for C.

$$C \sim IW\left(2, I_2\right).$$

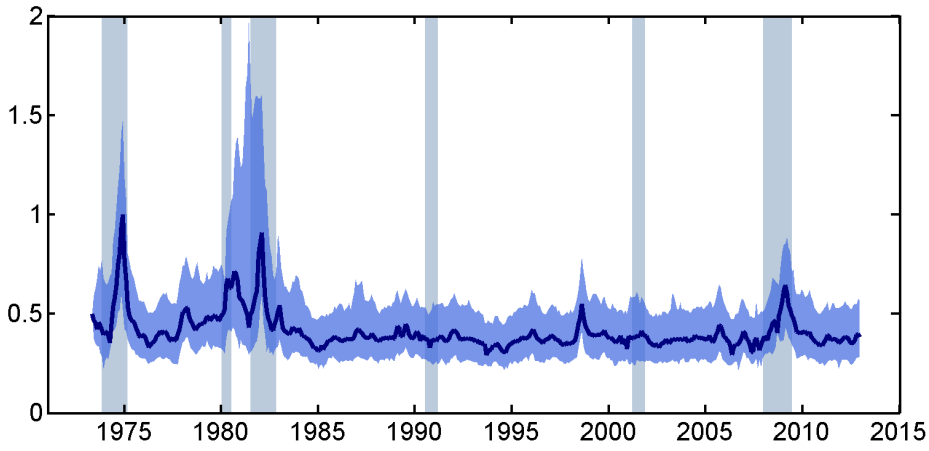
Finally, ν, the degrees of freedom parameter is set to be 50.

2.6.3. Posterior estimates

Stochastic volatility estimates

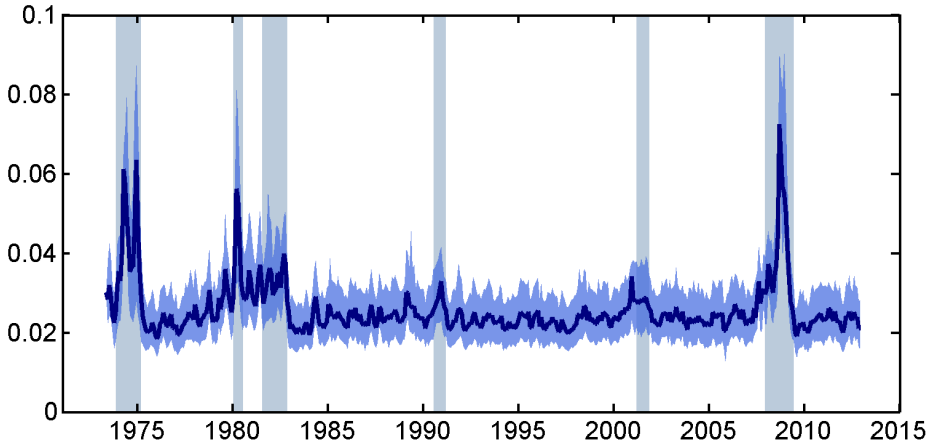
Estimated stochastic volatility series are presented in Figure 26 and 27.

Figure 26 Estimated Stochastic Volatility, IP growth



Note: Estimated stochastic volatility for IP growth based on CAIW(1)-in-VAR(4) model.

Figure 27 Estimated Stochastic Volatility, Financial conditions index



Note: Estimated stochastic volatility for Financial conditions index based on CAIW(1)-in-VAR(4) model.

Parameter estimates

Posterior estimates are presented in Table 21.

Table 21 Posterior Estimates for CAIW(1)-in-VAR(4)

Conditional mean equation			Volatility equation		
Parameter	Mean	90% Interval	Parameter	Mean	90% Interval
μ_1	-1.95	[-4.222, 0.007]	A_{11}	0.83	[0.776, 0.883]
μ_2	0.43	[0.123, 0.754]	A_{12}	0.89	[0.588, 1.197]
$\Phi_{11,1}$	0.08	[-0.018, 0.171]	A_{21}	-0.04	[-0.054, -0.015]
$\Phi_{12,1}$	0.14	[-0.227, 0.513]	A_{22}	0.74	[0.667, 0.796]
$\Phi_{21,1}$	0.02	[0.001, 0.039]	C_{11}	0.09	[0.059, 0.122]
$\Phi_{22,1}$	1.81	[1.711, 1.912]	C_{12}	0.00	[-0.003, 0.007]
$\Phi_{11,2}$	0.21	[0.137, 0.293]	C_{22}	0.01	[0.010, 0.016]
$\Phi_{12,2}$	-0.92	[-1.586, -0.235]			
$\Phi_{21,2}$	0.02	[0.002, 0.039]			
$\Phi_{22,2}$	-1.37	[-1.547, -1.193]			
$\Phi_{11,3}$	0.16	[0.086, 0.240]			
$\Phi_{12,3}$	0.66	[-0.032, 1.333]			
$\Phi_{21,3}$	-0.00	[-0.018, 0.017]			
$\Phi_{22,3}$	0.64	[0.480, 0.814]			
$\Phi_{11,4}$	-0.00	[-0.074, 0.073]			
$\Phi_{12,4}$	0.01	[-0.338, 0.365]			
$\Phi_{21,4}$	-0.00	[-0.022, 0.013]			
$\Phi_{22,4}$	-0.13	[-0.213, -0.041]			
B_{11}	-0.04	[-0.433, 0.331]			
B_{12}	-0.56	[-1.117, -0.047]			
B_{21}	0.00	[-0.066, 0.073]			
B_{22}	0.12	[0.039, 0.202]			

Note: For 90% credible interval we report the 5th and 95th percentile of the posterior distribution. We denote $\Phi_{ij,l}$ as the (i, j) element of Φ on the lth lag.

Assumption Procedure Corollary Theorem Definition Lemma

CHAPTER 3 : Does Realized Volatility Help Bond Yield Density Prediction?

3.1. Introduction

Researchers have found that time-varying volatility exists in U.S. government bond yields. In this paper, we introduce volatility proxy data in the hopes of better capturing this time-varying volatility for predictive purposes. To do so, we develop a general estimation approach to incorporate volatility proxy information into dynamic factor models with stochastic volatility. We apply it to the dynamic Nelson-Siegel (DNS) model of bond yields. We find that the higher frequency movements of the yields in the realized volatility data contain valuable information for the stochastic volatility and lead to significantly better density predictions, especially in the short term.

Our approach covers the existing classes of dynamic factor models with stochastic volatility. Specifically, we can account for stochastic volatility on the latent factors or stochastic volatility on the measurement errors. We derive a measurement equation to link realized volatility to the model-implied conditional volatility of the original observables. Incorporating realized volatility improves estimation of the stochastic volatility by injecting precise volatility information into the model.

The DNS model is a dynamic factor model that uses latent level, slope, and curvature factors to drive the intertemporal movements of the yield curve. This reduces the high-dimensional yields to be driven by just three factors. The level of the yield curve has traditionally been linked to inflation expectations while the slope to the real economy. Our

preferred specification introduces stochastic volatility on these latent factors. This leads to a nice interpretation of the stochastic volatility as capturing the uncertainty surrounding well-understood aspects of the yield curve. It also reduces the dimension of modeling the time-varying volatility of the yield curve.

We then compare this specification to several others in the DNS framework, including random walk dynamics for the factors and stochastic volatilities, as well as stochastic volatility on the yield measurement equation. In a forecasting horserace on U.S. bond yields, our preferred specification features slight improvements in the point forecast performance and significant gains in the density forecast performance. We also find that allowing for time-varying volatility is important for density prediction, especially in the short run. Unlike conditional mean dynamics, modeling volatility as first-order autoregressive processes rather than random walks leads to better predictive performance. Furthermore, having stochastic volatility on the factor equation better captures the time-varying volatility in the bond yield data when compared to stochastic volatility on the measurement equation. Finally, in addition to the standard forecast evaluation criterion, we also evaluate our models along economically meaningful dimensions in the forms of forecasting empirical level, slope, and curvature.

Our paper relates to the literature in three main areas.

First, our paper relates to work started by Barndorff-Nielsen and Shephard (2002) in incorporating realized volatility in models with time-varying volatility. Takahashi, Omori, and Watanabe (2009) use daily stock return data in combination with high-frequency realized volatility to more accurately estimate the stochastic volatility. Maheu and McCurdy (2011) show that adding realized volatility directly into a model of stock returns can improve density forecasts over a model that only uses level data, such as the EGARCH. Jin and Maheu (2013) propose a model of stock returns and realized covariance based on time-varying Wishart distributions and find that their model provides superior density forecasts for returns. There also exists work adding realized volatility in observation-driven volatil-

ity models (Shephard and Sheppard (2010) and Hansen, Huang, and Shek (2012)). As opposed to the other papers, we consider a dynamic factor model with stochastic volatility on the factor equation and use the realized volatility to help in the extraction of this stochastic volatility. In this sense, we bring the factor structure in the conditional mean to the conditional volatility as well. Cieslak and Povala (2013) have a similar framework in a no-arbitrage term structure model. Furthermore, we are the first paper to investigate the implications of realized volatility on bond yield density predictability.

Second, we contribute to a large literature on bond yield forecasting. Most of the work has been done on point prediction (see Diebold and Rudebusch (2012) and Duffee (2012) for excellent surveys). There has been, however, a growing interest in density forecasting. Egorov, Hong, and Li (2006) were the first to evaluate the joint density prediction performance of yield curve models. They overturn the point forecasting result of the superiority in random walk forecasts and find that affine term structure models perform better when forecasting the entire density, especially on the conditional variance and kurtosis. Hautsch and Ou (2012) and Hautsch and Yang (2012) add stochastic volatility to the DNS model by considering an independent AR(1) specification for the log volatilities of the latent factors. They do not do formal density prediction evaluation of the model, but give suggestive results of the possible improvements in allowing for time-varying volatility. Carriero, Clark, and Marcellino (2013) find that using priors from a Gaussian no-arbitrage model in the context of a VAR with stochastic volatility improves short-run density forecasting performance. Building on this previous work, we introduce potentially highly accurate volatility information into the model in the form of realized volatility. We also extend the forecast evaluation by considering joint density forecasting in the forms of empirical level, slope, and curvature.

Finally, we also add to a growing literature on including realized volatility information in bond yield models. Andersen and Benzoni (2010) and Christensen, Lopez, and Rudebusch (2011) view realized volatility as a benchmark on which to compare the fits of affine term

structure models. Cieslak and Povala (2013) are interested in using realized covariance to better extract stochastic volatility and linking the stochastic volatility to macroeconomic and liquidity factors. These papers focus on in-sample investigations of incorporating realized volatility in bond yield models. Our paper, on the other hand, considers the improvement from using realized volatility in out-of-sample point and density prediction.

In section 2, we introduce our methodology for incorporating volatility proxies into dynamic factor models in the context of the DNS model and other competitor specifications. We discuss the data in section 3. We present our estimation and forecast evaluation methodology in section 4. In section 5, we present in-sample and out-of-sample results. We conclude in section 6.

3.2. Model

We introduce the dynamic Nelson-Siegel model with stochastic volatility (DNS-SV) proposed by Hautsch and Ou (2012), Hautsch and Yang (2012), and Bianchi, Mumtaz, and Surico (2009). Then, we discuss the incorporation of realized volatility information into this framework. Finally, we consider alternatives to our main approach.

3.2.1. The Dynamic Nelson-Siegel model and time-varying bond yield volatility

Denote $y_t(\tau)$ as the continuously compounded yield to maturity on a zero coupon bond with maturity of τ periods at time t. Following Diebold and Li (2006), we consider the factor model for the yield curve,

$$y_t(\tau) = f_{l,t} + f_{s,t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + f_{c,t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \epsilon_t(\tau), \quad \epsilon_t \sim N(0,\, Q) \qquad (3.1)$$

where $f_{l,t}, f_{s,t}$ and $f_{c,t}$ serve as latent factors and ϵ_t is a vector that collects idiosyncratic component $\epsilon_t(\tau)$ for all maturities. As is well documented in the literature, the first factor

mimics the level of the yield curve, the second the slope, and the third the curvature. While some authors have estimated the λ, we fix it at 0.0609, noting from Diebold and Li (2006) and others that the value does not move around too much across time and that its estimation does not seem to affect the results. We assume that the Q matrix is diagonal. This leads to the natural interpretation of a few common factors driving the comovements in a large number of yields. All of the other movements in the yields are considered idiosyncratic. As suggested by Diebold and Li (2006), we model the dynamic factors as independent univariate first-order autoregressive processes, given by,

$$f_{i,t} = \mu_{f,i}(1 - \phi_{f,i}) + \phi_{f,i} f_{i,t-1} + \eta_{i,t}, \quad \eta_t \sim N(0,\ H_t) \tag{3.2}$$

for $i = l, s, c$. η_t is a vector that collects the innovations to each factor and its variance-covariance matrix H_t potentially varies over time. We also assume that idiosyncratic shocks ϵ_t and factor shocks η_t are independent. The independent factor specification of the setup means that the movements of the factors are unrelated to each other. While this may seem as a tight restriction at first blush, Diebold and Rudebusch (2012) point out that the assumption does not seem poor in so far as the factors are related to the principal components of the yield curve. Following Hautsch and Ou (2012), Hautsch and Yang (2012) and Bianchi, Mumtaz, and Surico (2009), we model the logarithm of the variance of the shocks to the factor equation as AR(1) processes,

$$h_{i,t} = \mu_{h,i}(1 - \phi_{h,i}) + \phi_{h,i} h_{i,t-1} + \nu_{i,t}, \quad \nu_{i,t} \sim N(0,\ \sigma^2_{h,i}) \tag{3.3}$$

for $i = l, s, c$. $\exp(h_{i,t})$ corresponds to the ith diagonal element of the variance-covariance matrix H_t. Since the conditional mean dynamics of the factors are specified as independent from each other, it makes sense to model the stochastic volatilities of the factor innovations as independent. We call this specification the DNS-SV model (dynamic Nelson-Siegel with stochastic volatility).

3.2.2. DNS-RV

We claim that by using high-frequency data to construct realized volatilities of the yields it is possible to aid in the extraction of the stochastic volatilities governing the level, slope, and curvature of the DNS-SV model. Using realized volatility to augment our algorithm should make estimation of the stochastic volatility parameters more accurate and produce a superior predictive distribution. Crucially, we need to find an appropriate linkage between our volatility proxy - realized volatility - and the stochastic volatility in the model. Given the definition of the model-implied conditional volatility, we propose

$$\begin{aligned} RV_t \approx Var_{t-1}(y_t) &= diag(\Lambda_f H_t \Lambda_f' + Q) \\ &= diag(\tilde{\Lambda}_f \tilde{H}_t \tilde{\Lambda}_f' + Q) \end{aligned} \tag{3.4}$$

where RV_t is the realized volatility of bond yields which has the same dimension as the bond yield vector y_t and Λ_f is factor loading matrix given by equation 3.1. We write the logarithm of volatility in deviation form $\tilde{h}_{i,t} = h_{i,t} - \mu_{h,i}$ for $i = l, s, c$. Then $\tilde{H}_t$ is a 3×3 diagonal matrix with each element corresponding to $e^{\tilde{h}_{i,t}}$ and $\tilde{\Lambda}_f = \Lambda_f[e^{\mu_l/2}, e^{\mu_s/2}, e^{\mu_c/2}]'$. The second equality in equation 3.4 comes from the change in notation. [1] Insofar as realized volatility provides an accurate approximation to the true underlying conditional time-varying volatility, equation 3.4 is the one that links this information to the model.

Upon adding measurement error, one can view equation 3.4 as a nonlinear measurement equation. In principle, we have several tools to handle this nonlinearity, including the particle filter. To keep estimation computationally feasible, we choose to take a first order Taylor approximation of the logarithm of this equation around $\tilde{h}_t = 0$ with respect to h_t. This leads to a set of linear measurement equations that links the realized volatility of the bond yields and the underlying factor volatility,

$$\log(RV_t) = \beta + \Lambda_h \tilde{h}_t + \zeta_t, \quad \zeta_t \sim N(0, S). \tag{3.5}$$

[1] This strategy of linking an observed volatility measure to the model is also used in other papers (Maheu and McCurdy (2011) in a univariate model and Cieslak and Povala (2013) in a multivariate context).

In the estimation, we add in measurement error ζ_t. We view equation 3.5 as an approximation to equation 3.4. We call this new model the dynamic Nelson-Siegel with realized volatility (DNS-RV) model. The difference between this model and DNS-SV comes from augmenting equation 3.1 with a new measurement equation 3.5. This equation has a constant β and a factor loading Λ_h. The parameter β comes from the linearization while we can interpret Λ_h as a loading for the factor volatility used to reduce the dimension of the volatility data, $\log(RV_t)$. This very naturally extends the dynamic factor model, which transforms high-dimensional data (y_t) into a few number of factors (f_t) via the factor loading matrix Λ_f. The volatility factor loading (Λ_h) is a function of other model parameters ($\Lambda_f,\ \mu_h$) with the functional form given by the linearized version of equation 3.4.[2] We can view this as a model-consistent restriction on the linkage between the conditional volatility of observed data, approximated by $\log(RV_t)$, and the factor volatility h_t. For the same reasons as in the baseline DNS-SV model, we set the S matrix to be diagonal. These are interpreted as idiosyncratic errors, and we therefore do not model them to be contemporaneously or serially correlated.

In summary, the DNS-RV model introduces a new measurement equation into the state space of the DNS-SV model.

(Measurement equation)

$$\begin{aligned} y_t &= \Lambda_f f_t + \epsilon_t, \quad \epsilon_t \sim N(0,\ Q) \\ \log(RV_t) &= \beta + \Lambda_h \tilde{h}_t + \zeta_t, \quad \zeta_t \sim N\left(0, S\right), \end{aligned} \tag{3.6}$$

(Transition equation)

$$\begin{aligned} f_{i,t} &= \mu_{i,f}(1-\phi_{i,f}) + \phi_{i,f} f_{i,t-1} + \eta_{i,t}, \quad \eta_{i,t} \sim N(0,\ h_{i,t}) \\ h_{i,t} &= \mu_{i,h}(1-\phi_{i,h}) + \phi_{i,h} h_{i,t-1} + \nu_{i,t}, \quad \nu_{i,t} \sim N(0,\ \sigma^2_{i,h}) \end{aligned} \tag{3.7}$$

[2]Detailed formulas for the volatility factor loading matrix can be found in the appendix.

Table 22 Model Specifications

Label	Factors (level, slope, curvature)	Conditional variance	Realized volatility
RW-C	Random walk	Constant	Not used
RW-SV	Random walk	log AR(1) in each factor	Not used
RW-RV	Random walk	log AR(1) in each factor	Used
DNS-C	Independent AR(1)	Constant	Not used
DNS-SV	Independent AR(1)	log AR(1) in each factor	Not used
DNS-RV	Independent AR(1)	log AR(1) in each factor	Used
RW-SV-RW	Random walk	Random walk	Not used
DNS-SV-RW	Independent AR(1)	Random walk	Not used
DNS-ME-SV	Independent AR(1)	log AR(1) in measurement equation	Not used
DNS-ME-RV	Independent AR(1)	log AR(1) in measurement equation	Used

Note: We list the specifications for the DNS model considered in this paper.

for $i = l, s, c$. In our application, both observed bond yields (y_t) and realized volatilities ($\log(RV)_t$) are 17×1 vectors. Moreover, both sets of variables have a factor structure with dynamics following the transition equations.

3.2.3. Alternative specifications

We have four classes of alternative specifications to compare forecasts to our baseline model. We briefly introduce them in this section and list all specifications considered in the paper in Table 22.

Dynamic Nelson-Siegel (DNS-C)

The first model is the standard Diebold-Li DNS model discussed at the beginning of the paper. It does not allow for stochastic volatility. This model has been shown to forecast the level of bond yields quite well, at times beating the random walk model of yields.

Dynamic Nelson-Siegel-Stochastic Volatility (DNS-SV)

The second is the DNS-SV model that adds stochastic volatility to the transition equation. It is summarized at the beginning of this section. By allowing for stochastic volatility, this model should improve upon the standard DNS model especially in the second moments, as

it can capture the time-varying volatility present in the bond yield data.

Dynamic Nelson-Siegel-Random Walk (RW)

Bond yield forecasting research (e.g. van Dijk, Koopman, van der Wel, and Wright (2013) and references therein) has shown that random walk specifications of the yield curve generally perform quite well. Oftentimes, the no-change forecast from a current period does best among a large group of forecasting models. It is in this sense that bond yield forecasting is difficult. Given these results, we also augment the DNS-C, DNS-SV, and DNS-RV model classes with random walk parameterizations of the factor processes.

Macroeconomic research (Cogley and Sargent (2005) and Justiniano and Primiceri (2008)) often specifies stochastic volatility as following a random walk. Doing so reduces the number of parameters estimated while also providing a simple no-change forecast benchmark for time-varying volatility. As long-horizon bond yield volatility links with macroeconomic volatility, we also have random walk specifications for the stochastic volatilities.

Dynamic Nelson-Siegel-Measurement Error Stochastic Volatility (+ Realized Volatility) (DNS-ME)

Koopman, Mallee, and Van der Wel (2010) argue that putting the time-varying conditional volatility on the measurement errors provides an improvement for the in-sample fit of the DNS class of models. To evaluate whether these results extend to forecasting as well, we model independent AR(1) specifications for the measurement error stochastic volatilities. While following this strategy greatly increases the number of parameters estimated, it could improve forecasting as each yield has its own stochastic volatility process. Therefore, as opposed to a time-varying H_t and constant Q matrix in the DNS-SV and DNS-RV setups, now H is constant while Q_t has stochastic volatility.

$$Q_t = \mathbf{Diagmat}(e^{\mathbf{h_t}}) \tag{3.8}$$

$$h_{i,t} - \mu_{h,i} = \phi_{h,i}(h_{i,t-1} - \mu_{h,i}) + \nu_{i,t}, \quad \nu_{i,t} \sim N(0,\ r_i), \tag{3.9}$$

for $i = 1, ..., N$ where N is the number of bond yields in the observation equation. $\mathbf{h_t}$ is a vector that collects all stochastic volatilities in measurement errors and $\mathbf{Diagmat}(\cdot)$ turns a vector into a diagonal matrix. Q_t remains a diagonal matrix, as equation 3.8 shows. We again model the logarithm of the variances as independent first order autoregressive processes.

We also consider incorporating realized volatility information into this model. Doing so leads to the following relationship

$$RV_t \approx Var_{t-1}(y_t) = diag(\Lambda_f H \Lambda_f' + Q_t). \tag{3.10}$$

As before, we do a first order Taylor approximation of the logarithm of this equation. We also add in a measurement error for estimation. However, in contrast to the DNS-RV model, we link each element in $\log(RV_t)$ with individual stochastic volatility $h_{i,t}$.

3.3. Data

We use a panel of unsmoothed Fama and Bliss (1987) U.S. government bond yields at the monthly frequency with maturities of 3, 6, 9, 12, 15, 18, 21 months and 2, 2.5, 3, 4, 5, 6, 7, 8, 9, 10 years from January 1981 to December 2009. This dataset is provided by Jungbacker, Koopman, and van der Wel (2013)[3]. To construct monthly realized volatility series, we use daily U.S. government bond yield data with the same maturities from January 2, 1981 to December 30, 2009 taken from the Federal Reserve Board of Governors with the

[3] <http://qed.econ.queensu.ca/jae/datasets/jungbacker001/>

methodology of Gürkaynak, Sack, and Wright (2007) [4] [5].

We construct the realized variance of each month's yields using daily bond yield data. The formula for realized variance at time t is

$$RV_t = \sum_{d=1}^{D} \left(\Delta y_{t+\frac{d}{D}} \right)^2 .$$

where D is the number of daily data in one time period t. This formula converges in probability to the true conditional variance as the sampling frequency goes to infinity under assumptions laid out in Andersen, Bollerslev, Diebold, and Labys (2003). Usually, there are around 21 days in each month, with less depending upon the number of holidays in a month that fall on normal trading days.

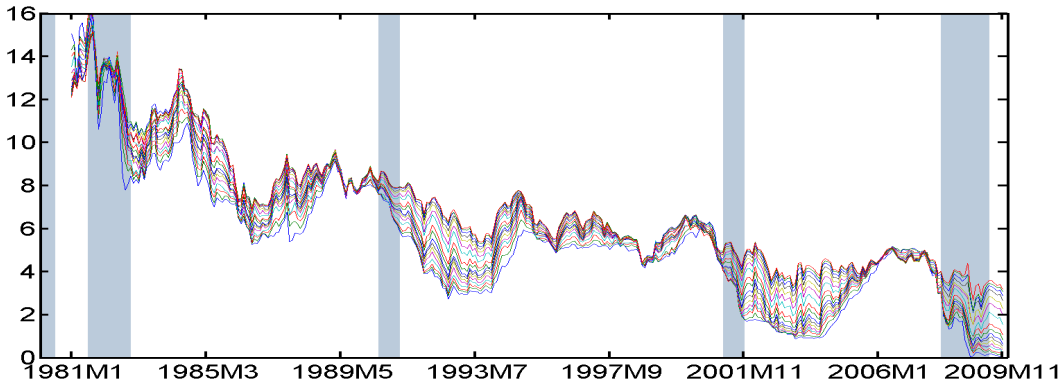
We use daily data to construct our realized volatilities for a few reasons. First, we want to use realized volatility information starting in 1981 to use a sample period similar to other bond yield forecasting studies. The availability of higher-frequency intraday data begins much later. For instance, Cieslak and Povala (2013) start their estimation in 1992 for specifically this reason. Second, the month-to-month volatility movements we want the volatility proxy to capture do not necessitate using ultra-high frequency data. Finally, while results may improve with higher frequency data, we show that positive effects are present even with lower frequency realized volatility.

[4] <http://www.federalreserve.gov/econresdata/researchdata/feds200628_1.html>

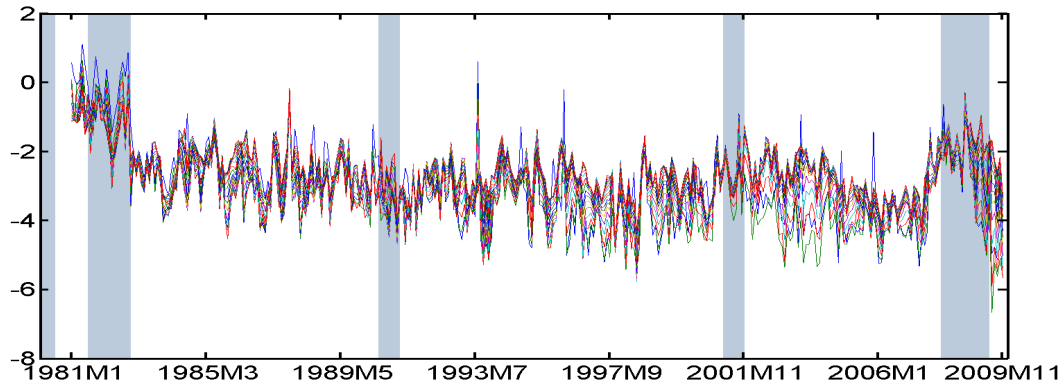
[5] Because most papers in the literature estimate the DNS model with unsmoothed Fama and Bliss data, we generate and evaluate predictions for the unsmoothed Fama and Bliss data. Unfortunately, this data is only available at the monthly frequency. Even though the two datasets use different methodologies, monthly yield data based on the daily bond yield from Gürkaynak, Sack, and Wright (2007) is very close to the one based on the unsmoothed Fama and Bliss method.

Figure 28 U.S. Treasury Yields

(a) Yields (monthly, annualized %)



(b) Yield Realized Volatilities (Monthly, Log)



Notes: We present monthly U.S. Treasury yields with maturities of 3, 6, 9, 12, 15, 18, 21 months and 2, 2.5, 3, 4, 5, 6, 7, 8, 9, 10 years over the period January 1981 – November 2009. Monthly yields are constructed using the unsmoothed Fama-Bliss method. Monthly yield realized volatilities are constructed based on daily yields using Wright's dataset. Blue shaded bars represent NBER recession dates.

Figure 28 plots the time series of monthly U.S. government bond yields and logarithm of realized volatilities. All yields exhibit a general downward trend from the start of the sample. For around the first 25 months, the realized volatility seems quite high and exhibits large time variation. After around 1983, yield volatility dies down and largely exhibits only temporary spikes in volatility. For a period of 2 years starting in 2008, the realized volatility picks up across all yields. We attribute this to the financial crises. Another interesting feature of log realized volatility is that it shows large autocorrelation. Its first-

Table 23 Variance explained by the first five principal components (%)

	Yield	log(RV)
pc 1	98.16	84.30
pc 2	99.84	94.62
pc 3	99.95	98.53
pc 4	99.98	99.46
pc 5	99.98	99.85

Notes: Numbers in the table are the percentage of total variance explained by the first five principal components for U.S. Yield data and log(RV). log(RV) is the logarithm of the monthly realized volatility constructed based on the daily U.S. yield data.

order autocorrelation coefficients range from 0.59 to 0.69 and the 12th-order autocorrelation coefficients range from 0.20 to 0.31.[6] This means that realized volatility data could help even for the long-horizon forecasts that we consider in this paper.

Both the yields and realized volatilities do seem to exhibit a factor structure, meaning that each set of series co-move over time. In fact, a principal components analysis shows that the first three factors for yields explain 99.95% of the variation in the U.S. yield curve. The first three factors for realized volatilities explain 98.53% of the variation (Table 23). Even though the fact that U.S. bond yield can be explained by the first few principal components is well documented, it is interesting that the same feature carries over for the realized volatility of U.S. bond yields.

3.4. Estimation/Evaluation Methodology

3.4.1. Estimation

We perform a Gibbs sampling Markov Chain Monte Carlo algorithm for 15000 draws. We keep every 5th draw and burn in for the first 5000 draws. Due to our linearization approximation in introducing realized volatility, all specifications that we consider can be sampled by using the method for the stochastic volatility state space model developed

[6] We provide tables in the supporting material for the descriptive statistics of monthly realized volatility of bond yields.

in Kim, Shephard, and Chib (1998). Details of the state space representation and the estimation procedure can be found in the appendix. Details on the prior can be found in the appendix as well, although we comment that our choice of prior is loose and does not impact the estimation results.

We highlight the difference in estimation procedure due to the additional measurement equation for the realized volatility. Roughly speaking, there are two sources of information for the latent volatility factor h_t. The first source is from the latent factor, f_t. To see this, one can transform the transition equation for f_t as the following.

$$\log\left((f_{i,t} - \mu_{i,f}(1-\phi_{i,f}) - \phi_{i,f} f_{i,t})^2\right) = h_{i,t} + \log(x_{i,t}^2), \quad x_{i,t} \sim N(0,1)$$

for $i = l, s, c$. This is common to both DNS-SV and DNS-RV. The second source is from equation 3.6 which relates $\log(RV_t)$ with h_t and is unique to DNS-RV. Following Kim, Shephard, and Chib (1998), we approximate $\log(x_{i,t}^2)$ with mixture of normals. Then, conditional on other parameters and data, extraction of h_t amounts to running a simulation smoother in conjunction with the Kalman filter with and without equation 3.6.

3.4.2. Forecast evaluation

We consider model performance along both the point and density forecasting dimensions. The appendix contains further details on the Bayesian simulation algorithm we use to generate the forecasts. We begin forecasting on February 1994 and reestimate the model in an expanding window and forecast moving forward two months at a time. For every forecast run at a given time t, we forecast for all yields in our dataset and for horizons ranging from 1 month to 12 months ahead. This leads to a total of 94 repetitions.

Point forecast

To evaluate the point prediction, we use the Root Mean Square Forecast Error (RMSE) statistic.

$$RMSE^{M}_{\tau,ho} = \sqrt{\frac{1}{F}\sum \left(\hat{y}^{M}_{t+ho}(\tau) - y(\tau)_{t+ho}\right)^2} \tag{3.11}$$

Call the yield τ forecast at horizon ho made by model M as $\hat{y}^{M}_{t+ho}(\tau)$ and $y(\tau)_{t+ho}$ the realized value of the yield at time $t + ho$. F is the number of forecasts made. Then, equation 3.11 provides the formula for the RMSE. To gauge whether there are significant differences in the RMSE, we use the Diebold and Mariano (1995) t-test of equal predictive accuracy.

Density forecast

The log predictive score (Geweke and Amisano (2010)) gives an indication of how well a model performs in density forecasting.

$$LPS^{M}_{\tau,ho} = \frac{1}{F}\sum \log p(y_{t+ho}(\tau)|y^t, M) \tag{3.12}$$

$p(y_{t+ho}(\tau)|y^t, M)$ denotes the ho-step ahead predictive distribution of yield τ generated by model M given time t information. Following Carriero, Clark, and Marcellino (2013), we estimate the log predictive density by a kernel density estimator using MCMC draws for parameters and latent states and compute the p-value for the Amisano and Giacomini (2007) t-test of equal means to gauge whether there exist significant differences in the log predictive score.

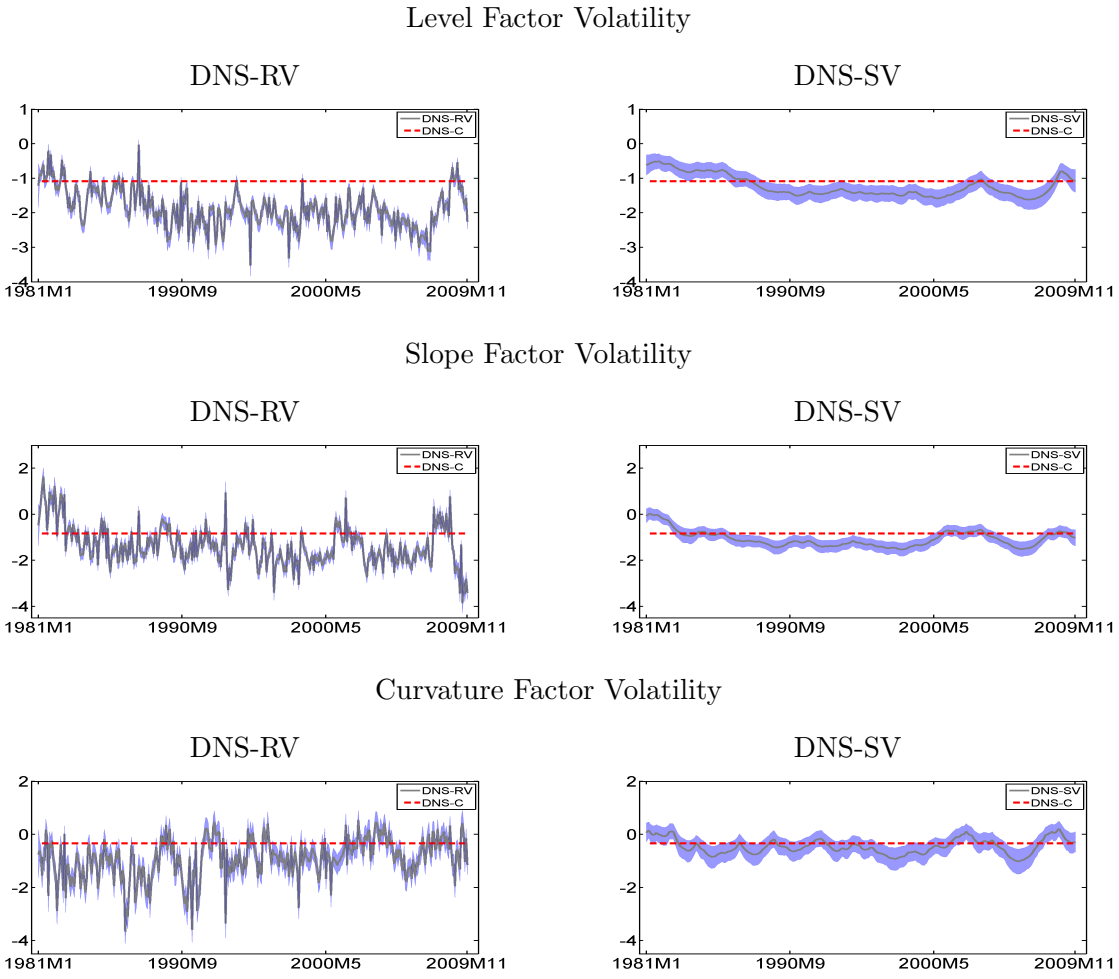
3.5. Results

We first present in-sample results of the model, focusing on time-varying volatility. Then, we move to point and density forecasting results. Finally, we discuss our empirical factor prediction exercise.

3.5.1. In-sample

We first present the full sample estimation from January 1981 - November 2009. We focus on how adding realized volatility information alters the model. Adding in second-moment information does not significantly change conditional mean dynamics, so we relegate our discussion of the extracted factors to the supporting material[7]. Our extracted factors are similar to those found in Diebold and Li (2006).

Figure 29 Stochastic volatility for bond yield factors



Notes: Posterior median of log stochastic volatility (h_t) for bond yield factors from DNS-RV (left column) and DNS-SV (right column). Red dotted line is volatility level estimated from DNS-C. Blue band is 80% credible interval. Estimation sample is from January 1981 to November 2009.

[7]All other in-sample estimation results are in the supporting material

Table 24 Posterior Estimates of Parameters on h_t Equation

	DNS-SV			DNS-RV		
	5%	50%	95%	5%	50%	95%
$\mu_{h,l}$	-4.40	-2.56	-1.48	-4.19	-3.92	-2.21
$\mu_{h,s}$	-3.41	-2.30	-1.67	-3.22	-2.85	-1.94
$\mu_{h,c}$	-1.48	-0.96	-0.47	-2.18	-1.88	-1.39
$\phi_{h,l}$	0.93	0.98	0.999	0.58	0.66	0.73
$\phi_{h,s}$	0.92	0.97	0.995	0.57	0.64	0.72
$\phi_{h,c}$	0.81	0.92	0.98	0.39	0.49	0.60
$\sigma_{h,l}$	0.01	0.03	0.08	0.41	0.75	0.92
$\sigma_{h,s}$	0.01	0.04	0.10	1.27	1.72	2.64
$\sigma_{h,c}$	0.02	0.09	0.27	1.30	1.78	4.30

Notes: Posterior moments are based on estimation sample from January 1981 to November 2009.

The stochastic volatility dynamics deserve some more precise discussion. Figure 29 shows the volatility estimate from the DNS-C, DNS-SV, and DNS-RV models. The fluctuations of the extracted stochastic volatilities in both the DNS-SV and DNS-RV models show that there exists conditional time-varying volatility in the data. Relative to the DNS-SV specification, adding in realized volatility data makes the extracted stochastic volatility much less persistent and more variable. This leads to a lower autoregressive parameter and higher innovation standard deviation estimates for all of the stochastic volatility processes (Table 24). The DNS-RV model also delivers lower stochastic volatility mean estimates. These differences lead to differences in forecasting. For example, the lower autoregressive parameter in the DNS-RV model means that it predicts faster mean reversion of the stochastic volatilities relative to the DNS-SV model and the lower long-run mean estimate implies that the DNS-RV model produces a tighter density prediction in the long run. The smoothed stochastic volatilities from the DNS-SV model, however, generally captures the low-frequency movements of those from the DNS-RV model.

We argue that this difference in the stochastic volatilities matters for density forecasting. The high-frequency data used to construct the realized volatilities brings information that the low-frequency monthly yield data misses. By having more accurate estimates of the

current level of time-varying volatility and process parameters, the DNS-RV model both starts off forecasting at a more accurate point and better captures the dynamics of the data moving forward.

3.5.2. Point prediction

Table 25 shows the RMSE of selected maturities for 1, 3, 6, and 12-step ahead predictions. The second column has the calculated RMSE values for the RW-C model. All other values reported are ratios relative to the RW-C RMSE. Values below 1 indicate superior performance relative to the random walk benchmark. Stars in the table indicate significant gains relative to the RW-C model. As expected, RMSE increases as the forecasting horizon lengthens. The models with random walk dynamics in the factors do well for short-horizon forecasts but deteriorate when compared to the stationary models as the prediction horizon lengthens. In general, all RMSE values have numbers close to 1, reproducing the well-known result in the bond yield forecasting literature on the difficulty in beating the no-change forecast. As alluded to in the previous section, adding in time-varying second moments does not largely impact point predictions, although the DNS-RV model forecasts middle maturities well across all horizons. For 12-month horizon forecasts, the DNS-C and DNS-ME-RV models also do well for short maturity yields.

3.5.3. Density prediction

Table 26 shows the density evaluation result in terms of log predictive score. Similar to the RMSE table, the RW-C column gives the value of the log predictive score for the random walk case while the numbers for the other models are differences relative to that column. A higher value indicates larger log predictive score and better density forecasting results. We present p-values based on Amisano and Giacomini (2007) comparing the hypothesis of equal log predictive score of the DNS-RV with those of the alternative models in table 27. As opposed to the point prediction results, three interesting findings emerge when we consider the log predictive score.

Table 25 RMSE comparison

Maturity	RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-SV-RW	DNS-SV-RW	DNS-ME-SV	DNS-ME-RV
1-step-ahead prediction										
3	0.267	0.992	1.011	1.004	1.034**	1.007	**0.991**	1.038**	1.071	0.953**
12	**0.229**	1.005	1.002	1.076**	1.055**	1.003	1.002	1.060**	1.070**	1.072**
36	0.274	1.001	0.996	1.015	1.020	**0.990**	1.001	1.025	1.010	1.012
60	0.274	1.000	0.995	1.003	1.010	**0.987**	1.000	1.013	1.002	1.004
120	0.277	1.001	0.996	0.988	0.989	0.999	1.000	0.988	**0.986**	0.997
3-step-ahead prediction										
3	**0.506**	1.002	1.011	1.036	1.066**	1.018	1.002	1.069**	1.055	1.024
12	0.537	1.002	**0.999**	1.072	1.073**	**0.999**	1.002	1.077**	1.064	1.063
36	0.580	1.000	0.998	1.022	1.036	**0.990**	0.999	1.040	1.018	1.018
60	0.551	1.000	0.998	1.004	1.017	**0.986**	0.999	1.021	1.005	1.004
120	0.489	1.000	0.998	**0.982**	0.987	0.995	1.000	0.988	0.987	0.987
6-step-ahead prediction										
3	0.932	1.000	1.002	1.003	1.047	1.007	**0.999**	1.048	1.010	1.000
12	0.915	1.001	1.001	1.040	1.065	**0.999**	1.000	1.068	1.036	1.034
36	0.881	1.001	1.001	1.017	1.041	**0.987**	1.000	1.046	1.011	1.011
60	0.819	1.000	0.998	0.998	1.016	**0.978**	1.000	1.020	0.997	0.997
120	0.665	0.999	0.992*	**0.977**	0.982	0.980	0.998	0.986	0.981	0.980
12-step-ahead prediction										
3	1.592	1.002*	0.999	**0.936****	1.013	0.991	1.000	1.015	0.943**	0.940*
12	1.476	1.001	0.999	**0.975**	1.034	0.991	1.000	1.036	0.976	**0.975**
36	1.242	1.001	1.001	0.997	1.043	**0.979****	1.001	1.046	0.993	0.993
60	1.069	1.002	1.000	0.997	1.031	**0.967***	1.001	1.033	0.994	0.994
120	0.833	1.002	0.998	0.976	0.994	**0.966**	1.001	0.994	0.980	0.976

Notes: The first column shows the RMSE based on the RW-C. Other columns show the relative RMSE compared to the first column. The RMSE from the best model for each variable and forecast horizon is in bold letter. Units are in percentage points. Divergences in accuracy that are statistically different from zero are given by * (10%), ** (5%), *** (1%). We construct the p-values based on the Diebold and Mariano (1995) t-statistics with a variance estimator robust to serial correlation using a rectangular kernel of $h-1$ lags and the small-sample correction proposed by Harvey, Leybourne, and Newbold (1997).

First, for the short-run horizon, having realized volatility gives significant gains in density prediction. This is on top of a large improvement in log predictive score from adding stochastic volatility, which Carriero, Clark, and Marcellino (2013) find. Table 27 shows that the DNS-RV model has significantly higher log predictive score values for one- and three-month ahead predictions for short term maturities when compared to most competitors.[8] By producing improved estimates of the current state of volatility, we would expect that short horizon forecasts have the largest gain. The improved density forecasting performance for the DNS-RV and RW-RV models continues even up to a 6-month forecasting horizon. At one year ahead, most models with realized volatility have their volatility processes returning close to the unconditional mean, so the gain diminishes.

[8]We also compute the model confidence set of Hansen, Lunde, and Nason (2011) and find the similar result. See the supporting material for the model confidence set.

Table 26 Log predictive score comparison

Maturity	RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-SV-RW	DNS-SV-RW	DNS-ME-SV	DNS-ME-RV
1-step-ahead prediction										
3	-0.538	0.205	0.343	0.008	0.198	**0.375**	0.181	0.173	0.048	0.102
12	-0.357	0.211	0.304	-0.009	0.195	**0.316**	0.189	0.172	0.012	0.013
36	-0.318	0.129	0.158	0.000	0.118	**0.176**	0.111	0.104	0.022	0.017
60	-0.265	**0.109**	0.077	-0.001	0.098	0.105	0.098	0.088	0.019	0.015
120	-0.248	**0.110**	0.111	0.011	**0.121**	0.110	0.106	0.108	0.048	0.037
3-step-ahead prediction										
3	-1.061	0.183	0.302	0.016	0.165	**0.322**	0.151	0.133	0.036	0.061
12	-0.985	0.107	0.195	-0.012	0.064	**0.196**	0.089	0.045	0.007	0.010
36	-0.951	0.027	0.053	-0.002	-0.002	**0.067**	0.013	-0.018	0.009	0.007
60	-0.882	0.013	0.016	0.003	-0.011	**0.035**	0.007	-0.015	0.011	0.010
120	-0.779	0.033	0.020	0.014	0.033	0.031	**0.036**	**0.038**	0.027	0.029
6-step-ahead prediction										
3	-1.485	0.048	0.157	0.028	0.023	**0.159**	0.033	0.005	0.038	0.053
12	-1.417	0.010	0.065	0.000	-0.058	**0.066**	0.000	-0.082	0.008	0.012
36	-1.342	-0.017	0.021	0.001	-0.077	**0.030**	-0.039	-0.095	0.010	0.012
60	-1.263	-0.031	0.008	0.010	-0.072	**0.018**	-0.045	-0.083	0.016	0.016
120	-1.100	0.031	0.049	0.023	0.038	**0.061**	0.029	0.033	0.036	0.034
12-step-ahead prediction										
3	-1.940	-0.096	0.031	**0.091**	-0.098	0.041	-0.096	-0.119	0.086	0.094
12	-1.850	-0.071	-0.012	**0.048**	-0.147	-0.004	-0.089	-0.168	0.048	0.051
36	-1.688	-0.017	-0.001	0.023	-0.122	0.019	-0.049	-0.137	0.032	**0.033**
60	-1.560	0.005	0.036	0.024	-0.068	**0.066**	-0.011	-0.073	0.034	0.036
120	-1.387	0.099	0.151	0.030	0.101	**0.193**	0.098	0.092	0.047	0.045

Notes: The first column shows the log predictive score based on the RW-C. Other columns show the difference of log predictive score from the first column. Log predictive score differences represent percentage point differences. Therefore, a difference of 0.1 corresponds to a 10% more accurate density forecast. The log predictive score from the best model is in bold letter for each variable and forecast horizon.

Second, comparing RW-SV to RW-SV-RW, RW-RV to RW-RV-RW, and DNS-SV to DNS-SV-RW shows that given a fixed conditional mean specification, a random walk specification on conditional volatility dynamics in general leads to poorer results. This illustrates the fact that even though conditional mean dynamics of bond yields approximate a random walk in our sample, conditional volatility dynamics exhibit mean reversion. Bond yields therefore do have forecastability, although simply looking at the conditional mean dynamics do not reveal this fact strongly.

Third, an alternative specification for introducing stochastic volatility into the model by putting it on the measurement equation does not forecast as well as the specification with stochastic volatility on the transition equation. The measurement error specifications give consistent improvements in the log predictive score over and above constant volatility models, although they have a similar log predictive score pattern as DNS-C. Comparing DNS-SV

Table 27 Log predictive score comparison: p-values

Maturity	RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	RW-SV-RW	DNS-SV-RW	DNS-ME-SV	DNS-ME-RV
1-step-ahead prediction									
3	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
12	**0.00**	**0.00**	0.13	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
36	**0.00**	0.22	**0.01**	**0.00**	0.18	0.10	0.11	**0.01**	**0.01**
60	0.08	0.93	**0.02**	0.10	0.88	0.87	0.72	0.17	0.15
120	0.07	0.99	0.94	0.11	0.82	0.93	0.98	0.30	0.22
3-step-ahead prediction									
3	**0.00**	**0.00**	0.17	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**
12	**0.00**	**0.03**	0.93	**0.00**	**0.01**	**0.02**	**0.00**	**0.00**	**0.00**
36	0.20	0.31	0.28	0.22	0.18	0.21	0.12	0.29	0.27
60	0.53	0.60	0.27	0.59	0.37	0.51	0.34	0.67	0.66
120	0.60	0.96	0.64	0.77	0.96	0.91	0.87	0.94	0.97
6-step-ahead prediction									
3	0.12	**0.04**	0.93	0.25	**0.04**	**0.02**	**0.03**	0.24	0.32
12	0.38	0.28	0.98	0.44	0.15	0.20	0.12	0.45	0.50
36	0.57	0.16	0.74	0.64	0.19	0.09	0.16	0.73	0.75
60	0.77	0.20	0.74	0.90	0.18	0.12	0.14	0.98	0.97
120	0.35	0.52	0.78	0.55	0.64	0.51	0.59	0.68	0.66
12-step-ahead prediction									
3	0.76	0.10	0.82	0.70	0.30	0.06	0.25	0.70	0.66
12	0.97	0.23	0.80	0.38	0.41	0.11	0.36	0.34	0.35
36	0.72	**0.02**	0.17	0.92	0.43	**0.00**	0.38	0.69	0.72
60	0.10	**0.00**	0.26	0.44	0.34	**0.00**	0.32	0.51	0.56
120	**0.01**	**0.00**	0.39	0.05	0.26	**0.00**	0.23	0.07	0.07

Notes: This table presents the p-values from Amisano and Giacomini (2007) tests comparing the hypothesis of equal log predictive score of the DNS-RV with alternative models. Bold letter indicates p-values less than 5%. Test statistics are computed with a variance estimator robust to serial correlation using a rectangular kernel of $h-1$ lags and the small-sample correction proposed by Harvey, Leybourne, and Newbold (1997).

to DNS-ME-SV shows that for short horizon forecasts, DNS-SV performs better whereas for longer horizon forecasts, DNS-ME-SV does better. A similar story holds when looking at DNS-RV and DNS-ME-RV, although DNS-RV does better even up to 6-month horizon forecasts with mixed 12-month horizon results.

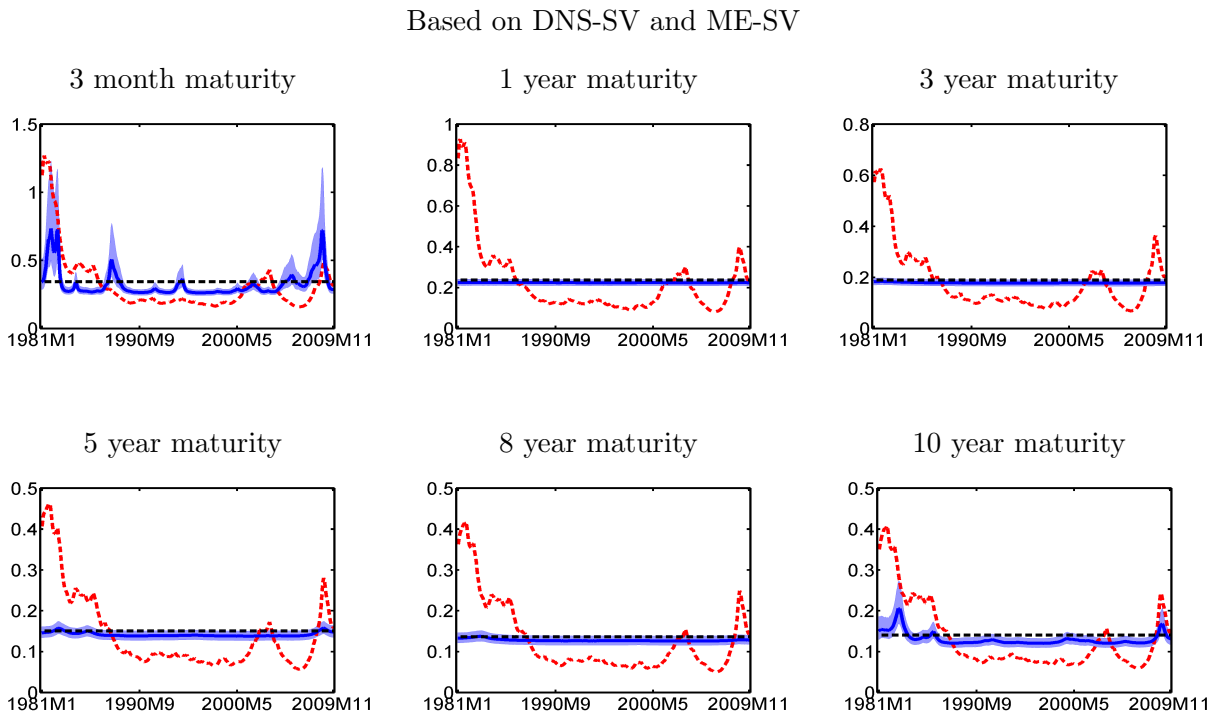
One fact holding back the performance of the measurement error specifications is that the measurement error variance explains a small portion of total bond yield variance. For example, the ratio of the standard deviations of smoothed measurement errors to the standard deviations of smoothed factors in the DNS-ME-SV model is often below 3 percent and never above 8 percent [9]. In figure 30, we see that the model with stochastic volatility on the measurement errors does not generate movements in the conditional time-varying volatility in

[9]This fact is similar across different model specifications. The values from all models are presented in the supporting materials.

various middle maturity yields. In fact, the conditional variances of the 1-year, 3-year, 5-year, and 8-year maturities are nearly on top of the black dotted line, which is the variance of the yields implied by the DNS-C model. Therefore, putting time-varying volatility in the measurement errors does not drastically change the model-implied predictive distributions. This explains why the density forecasting performance for this class of models mimics that of the DNS-C model.

In contrast, the figure 30 shows that putting stochastic volatility in the transition equation can better capture time-varying volatility. The DNS-SV model-implied time-varying volatility consistently fits the narrative evidence of the Great Moderation from the mid-1980's until the mid-2000's. It also picks up the increases in volatility from the early 2000's recession and recent financial crises.

Figure 30 Stochastic variance of individual yield



Notes: Red dotted line: Conditional variance from DNS-SV. Blue solid line: Conditional variance from DNS-ME-SV with 80% credible interval. Black dotted horizontal line: Conditional variance from DNS-C. Estimation sample is from January 1981 to November 2009.

3.5.4. Forecasting empirical factors

A popular decomposition that we use as well to model the yield curve involves the level, slope, and curvature factors. The empirical counterparts of these factors are the 10-year yield for the level, 10-year minus 6-month spread for the slope, and 6-month + 10-year − 2∗5-year linear combination for the curvature. As discussed in Diebold and Rudebusch (2012), the level factor is related to inflation expectations while the slope factor corresponds to the business cycle.

Given that many people think about the yield curve in terms of these three factors, we also provide point and density forecasting results in this dimension. Considering this dimension provides a unique challenge for our models, as they must jointly predict the future movements of the yields correctly. This is an aspect of multivariate forecasting not present in a univariate context.

Table 28 shows the RMSE values relative to the RW-C benchmark. In the short-run, RW-RV does well for slope and curvature and DNS-RV does well for curvature. All values, however, are not significantly different from a random walk, largely replicating the result for the individual yields. As the forecasting horizon increases, the stationary models begin to forecast well, as also pointed out in Duffee (2011). This is in line with results we find for the stationary models when forecasting the individual yields.

For the short- and long-run forecasts, allowing for time-varying volatility or adding realized volatility does not help in point prediction. Note that from table 25, adding in realized volatility largely improves the point prediction results when comparing DNS-SV to DNS-RV. This fact does not hold when forecasting the empirical factors. We believe there may be two factors at play that lead to this result. First, DNS-RV occasionally underperforms DNS-SV in forecasting the 10-year yield, an important component in the empirical factors. Second, joint prediction brings in the importance of the covariance between the yields. The root mean square error of the slope forecast, for example, depends on the RMSEs of the

6-month and 10-year yields, but also on the covariance between the errors of the two series. Specifically, a negative covariance between the forecast errors of the two series worsens RMSE. A negatively correlated forecast error could also cause the negative DNS-RV result. Perhaps by modeling the covariance between the yields more seriously and adding realized covariance data, second moment information could improve point forecasts.

For the density forecasts, allowing for stochastic volatility helps improve the joint prediction, especially in the short-run. Its effect diminishes as the horizon gets longer, similar to the individual yield results. In contrast to the individual yield prediction results, the random walk assumption on stochastic volatility no longer hurts prediction. Incorporating realized volatility information, however, worsens the density prediction of slope and curvature, although the difference is not significant for slope. A potential explanation for this result is that by only adding realized volatility information, we weight the DNS-RV estimates of h_t too strongly towards matching the individual variances of the yields, while neglecting the covariances. Adding realized covariance here could also reverse this fact.

Table 28 RMSE comparison: Empirical factors

Factor	RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-SV-RW	DNS-SV-RW	DNS-ME-SV	DNS-ME-RV
1-step-ahead prediction										
Level	0.277	1.001	**0.996**	0.988	0.989	0.999	1.000	0.988	0.986	0.997
Slope	0.295	0.995	**0.984**	1.014	1.008	1.007	0.998	1.010	1.020	0.999
Curvature	0.263	1.001	**0.968**	1.018	0.995	0.983	1.004	0.995	0.991	1.034
3-step-ahead prediction										
Level	0.489	1.000	0.998	**0.982**	0.987	0.995	1.000	0.988	0.987	0.987
Slope	0.489	0.999	**0.995**	1.023	1.018	1.027	1.001	1.018	1.023	1.020
Curvature	0.438	1.003	0.999	0.982	**0.955***	0.997	1.003	0.955*	0.974	0.996
6-step-ahead prediction										
Level	0.665	0.999	0.992*	**0.977**	0.982	0.980	0.998	0.986	0.981	0.980
Slope	0.773	1.000	0.996	**0.989**	1.012	1.034	1.000	1.012	0.990	0.991
Curvature	0.651	1.002	1.004	0.930	0.915***	0.995	1.002	**0.915*** **	0.930	0.942
12-step-ahead prediction										
Level	0.833	1.002	0.998	0.976	0.994	0.966	1.001	0.994	0.980	**0.976**
Slope	1.254	1.000	0.998**	**0.917**	0.985	1.013	1.000	0.986	0.920**	0.922**
Curvature	0.827	1.002**	1.003	0.908	0.923*	0.969	1.002***	0.922*	**0.905**	0.918

Notes: The first column shows the RMSE based on RW-C. Other columns show the relative RMSE compared to the first column. The RMSE from the best model is in bold letter. Units are in percentage points. Empirical factor is defined as: Level: 10-year yield, Slope: 10Y-6M spread, Curvature: 6M+10Y-2*5Y. Divergences in accuracy that are statistically different from zero are given by * (10%), ** (5%), *** (1%). We construct the p-values based on the Diebold and Mariano (1995) t-statistics with a variance estimator robust to serial correlation using a rectangular kernel of $h-1$ lags and the small-sample correction proposed by Harvey, Leybourne, and Newbold (1997).

Table 29 Log predictive score: Empirical factors

Factor	RW-C	RW-SV	RW-RV	DNS-C	DNS-SV	DNS-RV	RW-RV-RW	DNS-SV-RW	DNS-ME-SV	DNS-ME-RV
1-step-ahead prediction										
Level	-0.248	0.110	0.111	0.011	**0.121**	0.110	0.106	0.108	0.048	0.037
Slope	-0.250	0.130	0.104*	-0.003	0.115	0.078	**0.134**	0.126	0.050	0.061
Curvature	-0.133	0.052**	0.008	-0.007	0.062**	0.005	0.056**	**0.065****	0.037	0.010
3-step-ahead prediction										
Level	-0.779	0.033	0.020	0.014	0.033	0.031	0.036	**0.038**	0.027	0.029
Slope	-0.744	**0.063**	0.053	-0.015	0.037	0.022	0.060	0.038	0.007	0.015
Curvature	-0.605	0.040*	0.003	0.023	0.098***	0.007	0.043**	**0.103*****	0.031	0.014
6-step-ahead prediction										
Level	-1.100	0.031	0.049	0.023	0.038	**0.061**	0.029	0.033	0.036	0.034
Slope	-1.180	-0.255	-0.064	**0.008**	-0.182	-0.101	-0.141	-0.175	0.006	0.002
Curvature	-0.995	-0.034	-0.006	0.069	0.066*	0.003	-0.031	**0.069****	0.061	0.045
12-step-ahead prediction										
Level	-1.387	0.099	0.151	0.030	0.101	**0.193**	0.098	0.092	0.047	0.045
Slope	-1.683	-0.630	-0.057	**0.044**	-0.687	-0.088	-0.289	-0.364	0.019	0.017
Curvature	-1.234	-0.062	-0.031	**0.077**	-0.034	0.021	-0.077	-0.032	0.077	0.061

Notes: For this table, we perform a one-sided test with the alternative hypothesis that the alternative model has larger log predictive score than DNS-RV. The first column shows the log predictive score based on RW-C. Other columns show the difference of log predictive score from the first column. The log predictive score from the best model is in bold letter. Empirical factor is defined as: Level: 10-year yield, Slope: 10Y-6M spread, Curvature: 6M+10Y-2*5Y. Log predictive score differences represent percentage point differences. Therefore, a difference of 0.1 corresponds to a 10% more accurate density forecast. Gains in accuracy that are statistically different from zero are given by * (10%), ** (5%), *** (1%). We construct the p-values based on the Diebold and Mariano (1995) t-statistics with a variance estimator robust to serial correlation using a rectangular kernel of $h-1$ lags and the small-sample correction proposed by Harvey, Leybourne, and Newbold (1997).

3.6. Conclusion

We investigate the effects of introducing realized volatility information on U.S. bond yield density forecasting. We compare the performance of our benchmark model DNS-RV with a variety of different models proposed in the literature and find that the DNS-RV model produces superior density forecasts, especially for the short-run. In addition to this, incorpo-

rating time-varying volatility in general improves density prediction, time-varying volatility is better modeled as a stationary process as opposed to a random walk, and time-varying volatility in the factor equation generates better density predictions when compared to time-varying volatility on the measurement equation.

The results for the joint forecasting performance show that there exists promising future work to be done in this area. Key in joint forecasting performance is capturing the correct time-varying covariance in the yields. Perhaps explicitly modeling correlation between the stochastic volatility innovations in conjunction with using realized covariance data could lead to large gains in joint density forecasting.

3.7. Appendix

3.7.1. State Space Representation

For completeness, we present the full specification of the state space form of the model. We give a detailed explanation of these equations in sections 2.1 and 2.2 of the main text. Consider a set of bond yields $y_t = \{y_t(1), ..., y_t(N)\}'$. τ_j is the maturity in months of bond yield j and λ is the point of maximal curvature.

$$y_t = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{\lambda\tau_1} \\ . & . & . \\ . & . & . \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{\lambda\tau_N} \end{pmatrix} \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix} + \epsilon_t \tag{3.13}$$

$$\epsilon_t \sim N(0, Q) \tag{3.14}$$

$$\log(RV_t) = \beta + \Lambda_h \tilde{h}_t + \zeta_t \tag{3.15}$$

$$\zeta_t \sim N(0, S) \tag{3.16}$$

$$\begin{pmatrix} l_t - \mu_l \\ s_t - \mu_s \\ c_t - \mu_c \end{pmatrix} = \begin{pmatrix} \phi_l & 0 & 0 \\ 0 & \phi_s & 0 \\ 0 & 0 & \phi_c \end{pmatrix} \begin{pmatrix} l_{t-1} - \mu_l \\ s_{t-1} - \mu_s \\ c_{t-1} - \mu_c \end{pmatrix} + \eta_t \tag{3.17}$$

$$\eta_t \sim N(0, H_t) \tag{3.18}$$

$$H_t = \begin{pmatrix} e^{h_{l,t}} & 0 & 0 \\ 0 & e^{h_{s,t}} & 0 \\ 0 & 0 & e^{h_{c,t}} \end{pmatrix} \tag{3.19}$$

$$h_{i,t} - \mu_{i,h} = \phi_{i,h}(h_{i,t-1} - \mu_{i,h}) + e_{i,t} \tag{3.20}$$

$$e_{i,t} \sim N(0, \sigma^2_{i,h}) \tag{3.21}$$

for $i = l, s, c$ and Q and S are diagonal matrix.

3.7.2. Measurement Equation for RV: Derivation and Approximation

Equation 3.15 is the linearized version of the nonlinear measurement equation that comes from adding realized volatility information to the dynamic factor model. We perform a first-order approximation of the logarithm of the following equation

$$\begin{aligned} RV_t \approx Var_{t-1}(y_t) &= diag(\Lambda_f H_t \Lambda_f' + Q) \\ &= diag(\tilde{\Lambda}_f \tilde{H}_t \tilde{\Lambda}_f' + Q) \end{aligned} \tag{3.22}$$

where we write the logarithm of volatility in deviation form $\tilde{h}_{i,t} = h_{i,t} - \mu_{h,i}$ for $i = l, s, c$. Then $\tilde{H}_t$ is a 3×3 diagonal matrix with each element corresponding to $e^{\tilde{h}_{i,t}}$ and $\tilde{\Lambda}_f = \Lambda_f[e^{\mu_l/2}, e^{\mu_s/2}, e^{\mu_c/2}]'$. We first derive the nonlinear measurement equation that links the realized volatility with underlying factor volatility. Our derivation is similar to Maheu and McCurdy (2011) but we derive it under the dynamic factor model framework. Then, we describe the approximation to get the linearized measurement equation for RV_t.

Derivation of the measurement equation We assume that the RV_t has a log-Normal distribution (element-wise). Then we have

$$E_{t-1}[RV_t] = \exp\left(E_{t-1}\log(RV_t) + \frac{1}{2}Var_{t-1}(\log(RV_t))\right) = diag(\tilde{\Lambda}_f \tilde{H}_t \tilde{\Lambda}_f' + Q).$$

where the second equality is from Corollary 1 of Andersen, Bollerslev, Diebold, and Labys (2003) and we assume that RV_t is an unbiased estimator for the quadratic variation. Taking logarithm on both sides gives,

$$E_{t-1}\log(RV_t) + \frac{1}{2}Var_{t-1}(\log(RV_t)) = \log\left(diag(\tilde{\Lambda}_f \tilde{H}_t \tilde{\Lambda}_f' + Q)\right).$$

Assume the conditional variance of the RV_t is constant and write $E_{t-1}[\log(RV_t)] = \log(RV_t) + \zeta_t$ where[10] $\zeta_t \sim \mathcal{N}(0, S)$. Then we get

$$\log(RV_t) = \tilde{\beta} + \log(diag(\tilde{\Lambda}_f \tilde{H}_t \tilde{\Lambda}_f' + Q)) + \zeta_t, \quad \zeta_t \sim \mathcal{N}(0, S). \tag{3.23}$$

where $\tilde{\beta}$ can be viewed as the conditional variance of log realized volatility plus potential bias caused by the assumption we made.

Linearization We present the derivation of equation 3.5. We linearize the equation 3.23 for ith element around $\tilde{h}_{j,t} = 0$ to get

$$\log(RV_{i,t}) = \beta_i + \nu_i \left(\sum_{j=1}^{3} \tilde{\Lambda}_{f,i,j}^2 \tilde{h}_{j,t}\right) + \zeta_t, \quad \zeta_t \sim \mathcal{N}(0, S).$$

where

$$\beta_i = \tilde{\beta}_i + \log(\tilde{\Lambda}_{f,i}^2 + Q_{i,i})$$
$$\nu_i = \frac{1}{\left(\tilde{\Lambda}_{f,i}^2 + Q_{i,i}\right)}.$$

[10] Distribution of ζ_t is obtained by assuming that RV_t follows a log-Normal distribution.

$\nu_i \left(\sum_{j=1}^{3} \tilde{\Lambda}^2_{f,i,j} \tilde{h}_{j,t} \right)$ corresponds to $\Lambda_{h,i} \tilde{h}_t$ and $\tilde{\Lambda}^2_{f,i,j}$ is the (i,j)th element of $\tilde{\Lambda}^2_f$ in equation 3.5. Note that $\beta \neq \tilde{\beta}$ as β is the constant term from the linearized equation while $\tilde{\beta}$ is from the nonlinear one.

3.7.3. Estimation Procedure

Presented is the algorithm for the Gibbs sampling. We draw 15000 samples, saving every 5th draw, with the first 5000 draws as burn-in. The priors we choose for the model are all extremely loose. They are presented in table 30.

Because of the assumption of independent AR(1) factor and stochastic volatility processes, the algorithm simplifies slightly. A general multivariate case, however, is a trivial extension.

Call $\Theta^* = \{\mu_f, \phi_f, \mu_h, \phi_h, \beta, Q, S, \sigma_h^2, f_{1:T}, h_{1:T}\}$, the parameters on which we would like to perform inference. Note that as is standard in Bayesian estimation, we use the data augmentation method and consider f_t and h_t as random vectors.

1. **(Initialize Θ^*)** We do so using the Hautsch and Yang (2012) estimates where possible. For S, we initialize with the identity matrix. For each element of β, we initialize at 0. For the factors and stochastic volatilities we draw each element from a normal distribution.

   Enter into iteration i:

2. **(Drawing $Q|\Theta^*_{-Q}$)** Since Q is diagonal, we draw the diagonal elements one at a time. Each element of the diagonal term on Q is distributed as an inverse gamma distribution.[11]

3. **(Drawing $\beta, S|\Theta^*_{-\beta,S}$)** We can likewise draw β and the diagonal elements of S equation-by-equation. It is a standard linear regression normal-inverse gamma framework.

[11] For DNS-RV, Q enters in the realized measurement equation (equation 3.15). In this case, we draw Q using the Metropolis-Hastings algorithm with a proposal distribution as an inverse gamma distribution.

4. **(Drawing $f_{1:T}|\Theta^*_{-f_{1:T}}$)** The Carter and Kohn (1994) multi-move Gibbs sampling procedure with stochastic volatility can be used to draw the level, slope, and curvature factors.

5. **(Drawing $\mu_f, \phi_f|\Theta^*_{-\mu_f,\phi_f}$)** Because we specify the factors and stochastic volatilities to have independent AR(1) processes, we can separate the drawing of the parameters for each factor. Drawing the parameters equation-by-equation is possible through the linear regression with stochastic volatility laid out in Hautsch and Yang (2012).

6. **(Drawing $h_{1:T}|\Theta^*_{-h_{1:T}}$)** We have a measurement equation made up of two parts. The first part uses the Kim, Shephard, and Chib (1998) method to transform the level, slope, and curvature factor equations. The measurement equation defined by the level factor is

$$\log((l_t - (1-\phi_{l,h})\mu_{l,h} - \phi_{l,h}l_{t-1})^2) = h_{l,t} + \log(x^2_{l,t}) \tag{3.24}$$

We approximate the error term using a mixture of normals as in Kim, Shephard, and Chib (1998).

The second part is the realized volatility measurement equation.

Therefore, equations 3.25 and 3.15 define the measurement equation:

$$f^*_t = \mu_h + I_{3\times3}(h_t - \mu_h) + \log(x^2_t) \tag{3.25}$$

The transition equation is

$$\begin{pmatrix} h_{l,t} - \mu_{l,h} \\ h_{s,t} - \mu_{s,h} \\ h_{c,t} - \mu_{c,h} \end{pmatrix} = \begin{pmatrix} \phi_{l,h} & 0 & 0 \\ 0 & \phi_{s,h} & 0 \\ 0 & 0 & \phi_{c,h} \end{pmatrix} \begin{pmatrix} h_{l,t-1} - \mu_{l,h} \\ h_{s,t-1} - \mu_{s,h} \\ h_{c,t-1} - \mu_{c,h} \end{pmatrix} + e_t \tag{3.26}$$

Because of our linear approximation of the nonlinear measurement equation, we can simply use the standard Kalman Filter along with the Carter and Kohn (1994) multi-move Gibbs simulation smoother with time-varying measurement mean and innovation volatility to draw $h_{1:T}$.

7. **(Drawing $\mu_h, \phi_h, \sigma_h^2 | \Theta^*_{-\mu_h,\phi_h,\sigma_h^2}$)** We use a standard linear regression normal-inverse gamma framework to draw the parameters equation-by-equation[12].

3.7.4. Forecasting Procedure

Presented in equations 3.27 - 3.29 is the forecasting algorithm that we use. Because we are performing Bayesian analysis, we explicitly take into account the parameter uncertainty when generating our forecasts. We first draw parameters from the relevant posterior distributions (j) and then simulate 10 trajectories of data given the parameter values (k). We do so for 2000 parameter draws for a total of 20000 simulated data chains from which to compare to the realized data (Del Negro and Schorfheide (2013)). Note that for the DNS-C model, we would not have equation 3.29 and the H_t would become H.

$$\hat{y}_t^{j,k} = \Lambda_f \begin{pmatrix} l_t^{j,k} \\ s_t^{j,k} \\ c_t^{j,k} \end{pmatrix} + \tilde{\epsilon}_t^{j,k}, \quad \tilde{\epsilon}_t^{j,k} \sim N\left(0, Q^j\right) \tag{3.27}$$

[12]Here relies on the specifics of the linearization for our realized volatility measurement equation. Note that we linearize that equation around the previous draw's μ_h. This means we do not have to take this equation into account when drawing our new μ_h.

$$\begin{pmatrix} l_t^{j,k} - \mu_l^j \\ s_t^{j,k} - \mu_s^j \\ c_t^{j,k} - \mu_c^j \end{pmatrix} = \begin{pmatrix} \phi_l^j & 0 & 0 \\ 0 & \phi_s^j & 0 \\ 0 & 0 & \phi_c^j \end{pmatrix} \begin{pmatrix} l_{t-1}^{j,k} - \mu_l^j \\ s_{t-1}^{j,k} - \mu_s^j \\ c_{t-1}^{j,k} - \mu_c^j \end{pmatrix} + \tilde{\eta}_t^{j,k}, \quad \tilde{\eta}_t^{j,k} \sim N\left(0, H_t^{j,k}\right) \tag{3.28}$$

$$h_{i,t}^{j,k} - \mu_{i,h}^j = \phi_{i,h}^j (h_{i,t-1}^{j,k} - \mu_{i,h}^j) + \tilde{e}_{i,t}^{j,k}, \quad \tilde{e}_{i,t}^{j,k} \sim N\left(0, \left(\sigma_{i,h}^j\right)^2\right) \tag{3.29}$$

$$j = 1, ..., 2000$$

$$k = 1, ..., 10$$

$$t = T, ..., T + 12$$

where T is the beginning of the forecasting period.

3.7.5. Prior Specification

We present prior distributions in table 30.

Table 30 Prior Distribution

Parameter	Description	Dim.	Dist.	Para(1)	Para(2)
H	Variance of the measurement error (y_t).	17×1	IG	0	0.001
μ_f	Long-run mean parameter for f_t.	3×1	N	0	100
ϕ_f	AR(1) coefficient for f_t.	3×1	N	0.8	100
μ_h	Long-run mean parameter for h_t.	3×1	N	0	100
ϕ_h	AR(1) coefficient for h_t.	3×1	N	0.8	100
σ_h^2	Variance of the innovation for the h_t.	3×1	IG	0.01	2
β	Intercepts in the RV measurement equation. Only used for $\mathcal{M}_{RV}$	17×1	N	0	100
S	Variance of the measurement error RV. Only used for $\mathcal{M}_{RV}$	17×1	IG	0	0.001
σ_f^2	Variance of the innovation for the f_t. Only used for models without time-varying volatility	3×1	IG	0.1	2

Note: a) All prior distributions are independent. For example, prior distributions for elements in H are independent from each other and follow the inverse gamma distribution.
b) Dim: Dimension of the parameters.
c) IG: Inverse gamma distribution. Para(1) and Para(2) mean scale and shape parameters, respectively.
d) N: Normal distribution. Para(1) and Para(2) stand for mean and variance, respectively.
e) Priors for ϕ_f and ϕ_h are truncated so that the processes for factors and volatilities are stationary.
f) $\mathcal{M}_{RV}$ is the set of models with realized volatility data.

BIBLIOGRAPHY

Alessandri, P., and H. Mumtaz (2014a): "Financial Conditions and Density Forecasts for US Output and Inflation," Working Papers, Queen Mary, University of London, School of Economics and Finance 715, Queen Mary, University of London, School of Economics and Finance.

——— (2014b): "Financial Regimes and Uncertainty Shocks," Working Papers 729, Queen Mary, University of London, School of Economics and Finance.

Amisano, G., and R. Giacomini (2007): "Comparing Density Forecasts via Weighted Likelihood Ratio Tests," *Journal of Business & Economic Statistics*, 25(2), 177–190.

Andersen, T. G., and L. Benzoni (2010): "Do Bonds Span Volatility Risk in the U.S. Treasury Market? A Specification Test for Affine Term Structure Models," *The Journal of Finance*, 65(2), 603–653.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71(2), 579–625.

Andreasen, M. M., J. Fernandez-Villaverde, and J. Rubio-Ramirez (2013): "The Pruned State-Space System for Non-Linear DSGE Models: Theory and Empirical Applications," NBER Working Papers 18983, National Bureau of Economic Research, Inc.

Angeletos, G.-M., F. Collard, and H. Dellas (2014): "Quantifying Confidence," Working papers.

Angeletos, G.-M., and J. La'O (2013): "Sentiments," *Econometrica*, 81(2), 739–779.

Aruoba, S. B., L. Bocola, and F. Schorfheide (2013a): "Assessing DSGE Model Nonlinearities," NBER Working Papers 19693, National Bureau of Economic Research, Inc.

Aruoba, S. B., L. Bocola, and F. Schorfheide (2013b): "Assessing DSGE Model Nonlinearities," Working Papers 13-47, Federal Reserve Bank of Philadelphia.

Aruoba, S. B., P. Cuba-Borda, and F. Schorfheide (2013): "Macroeconomic Dynamics Near the ZLB: A Tale of Two Countries," NBER Working Papers 19248, National Bureau of Economic Research, Inc.

Atchadé, Y. F., and J. S. Rosenthal (2005): "On Adaptive Markov Chain Monte Carlo Algorithms," *Bernoulli*, 11(5), 815–828.

Baker, S., N. Bloom, and S. Davis (2013): "Measuring Economic Policy Uncertainty," Discussion paper, policyuncertainty.com.

Bansal, R., and A. Yaron (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59(4), 1481–1509.

Barndorff-Nielsen, O. E., and N. Shephard (2002): "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(2), 253–280.

Barsky, R. B., and E. R. Sims (2012): "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence," *American Economic Review*, 102(4), 1343–77.

Beaudry, P., and F. Portier (2004): "An exploration into Pigou's theory of cycles," *Journal of Monetary Economics*, 51(6), 1183–1216.

——— (2006): "Stock Prices, News, and Economic Fluctuations," *American Economic Review, American Economic Association*, 96(4), 1293–1307.

Bianchi, F. (2013): "Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics," *The Review of Economic Studies*, 80(2), 463–490.

Bianchi, F., and L. Melosi (2013): "Modeling the Evolution of Expectations and Uncertainty in General Equilibrium," Discussion paper.

Bianchi, F., H. Mumtaz, and P. Surico (2009): "The Great Moderation of the Term Structure of UK Interest Rates," *Journal of Monetary Economics*, 56(6), 856 – 871.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013): "News, Noise, and Fluctuations: An Empirical Exploration," *American Economic Review*, 103(7), 3045–70.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.

Caldara, D., C. Fuentes-Albero, S. Gilchrist, and E. Zakrajsek (2013): "On the Identification of Financial and Uncertainty Shocks," 2013 Meeting Papers 965, Society for Economic Dynamics.

Carriero, A., T. E. Clark, and M. Marcellino (2013): "No Arbitrage Priors, Drifting Volatilities, and the Term Structure of Interest Rates," Discussion paper, Working paper.

Carter, C. K., and R. Kohn (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81(3), 541–553.

Christensen, J. H., J. A. Lopez, and G. D. Rudebusch (2011): "Can Spanned Term Structure Factors Drive Stochastic Volatility?," Discussion paper, Working paper.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

Christiano, L. J., R. Motto, and M. Rostagno (2014): "Risk Shocks," *American Economic Review*, 104(1), 27–65.

Cieslak, A., and P. Povala (2013): "Information in the Term Structure of Yield Curve Volatility," Discussion paper, Working paper.

Clark, T. E. (2011): "Real-Time Density Forecasts from Bayesian Vector Autoregressions with Stochastic Volatility," *Journal of Business & Economic Statistics*, 29(3).

Cogley, T., and T. J. Sargent (2005): "Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S," *Review of Economic Dynamics*, 8(2), 262–302.

Comin, D., and M. Gertler (2006): "Medium-Term Business Cycles," *American Economic Review*, 96(3), 523–551.

Croce, M. M. (2014): "Long-run productivity risk: A new hope for production-based asset pricing?," *Journal of Monetary Economics*, 66(0), 13 – 31.

Davis, J. (2007): "News and the Term Structure in General Equilibrium," Unpublished manuscript.

Davis, S. J., and J. Haltiwanger (2001): "Sectoral job creation and destruction responses to oil price changes," *Journal of Monetary Economics*, 48(3), 465–512.

Del Negro, M., and F. Schorfheide (2012): "DSGE model-based forecasting," Discussion paper.

——— (2013): "DSGE Model-Based Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott, and A. Timmermann, vol. 2A, chap. 2, pp. 57 – 140. Elsevier.

Diebold, F., J. Lee, and G. Weinbach (1994): "Regime Switching with Time-Varying Transition Probabilities," *Nonstationary Time Series Analysis and Cointegration*, pp. 238–302.

Diebold, F. X., and C. Li (2006): "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, 130(2), 337 – 364.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 20(1), 134–144.

Diebold, F. X., and R. S. Mariano (2002): "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics, American Statistical Association*, 20(1), 134–44.

Diebold, F. X., and G. D. Rudebusch (2012): *Yield Curve Modeling and Forecasting*. Princeton University Press.

Duffee, G. R. (2011): "Forecasting with the Term Structure: The Role of No-Arbitrage Restrictions," Discussion paper, Working papers//the Johns Hopkins University, Department of Economics.

Duffee, G. R. (2012): "Forecasting Interest Rates," Economics Working Paper Archive 599, The Johns Hopkins University,Department of Economics.

Egorov, A. V., Y. Hong, and H. Li (2006): "Validating Forecasts of the Joint Probability Density of Bond Yields: Can Affine Models Beat Random Walk?," *Journal of Econometrics*, 135(1 – 2), 255 – 284.

Elder, J. (2004): "Another Perspective on the Effects of Inflation Uncertainty," *Journal of Money, Credit and Banking*, 36(5), 911–28.

Elder, J., and A. Serletis (2010): "Oil Price Uncertainty," *Journal of Money, Credit and Banking*, 42(6), 1137–1159.

Fernald, J. (2012): "A Quarterly, Utilization-adjusted Series on Total Factor Productivity," Working Paper Series 2012-19, Federal Reserve Bank of San Francisco.

Fernandez-Villaverde, J., G. Gordon, P. A. Guerron-Quintana, and J. Rubio-Ramirez (2012): "Nonlinear Adventures at the Zero Lower Bound," NBER Working Papers 18058, National Bureau of Economic Research, Inc.

Fernandez-Villaverde, J., P. Guerron-Quintana, K. Kuester, and J. Rubio-Ramirez (2013): "Fiscal Volatility Shocks and Economic Activity," PIER Working Pa-

per Archive 11-022, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Fernandez-Villaverde, J., P. Guerron-Quintana, J. F. Rubio-Ramirez, and M. Uribe (2011): "Risk Matters: The Real Effects of Volatility Shocks," *American Economic Review*, 101(6), 2530–61.

Fernandez-Villaverde, J., and J. F. Rubio-Ramirez (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74(4), 1059–1087.

Filardo, A. J. (1994): "Business-Cycle Phases and Their Transitional Dynamics," *Journal of Business & Economic Statistics*, 12(3), 299–308.

Foerster, A., J. Rubio-Ramirez, D. Waggoner, and T. Zha (2013): "Perturbation methods for Markov-switching DSGE model," Discussion paper.

Fox, E. B., and M. West (2013): "Autoregressive Models for Variance Matrices: Stationary Inverse Wishart Processes," Discussion paper.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987): "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 19(1), 3–29.

Fujiwara, I., Y. Hirose, and M. Shintani (2011): "Can News Be a Major Source of Aggregate Fluctuations? A Bayesian DSGE Approach," *Journal of Money, Credit and Banking*, 43(1), 1–29.

Gertler, M., and N. Kiyotaki (2010): "Financial Intermediation and Credit Policy in Business Cycle Analysis," in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3 of *Handbook of Monetary Economics*, chap. 11, pp. 547–599. Elsevier.

Geweke, J., and G. Amisano (2010): "Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns," *International Journal of Forecasting*, 26(2), 216 – 230.

Giannone, D., M. Lenza, and G. E. Primiceri (2012): "Prior Selection for Vector Autoregressions," Working Papers ECARES ECARES 2012-002, ULB – Universite Libre de Bruxelles.

Giglio, S., B. T. Kelly, and S. Pruitt (2013): "Systemic Risk and the Macroeconomy: An Empirical Evaluation," Chicago booth research paper no. 12-49; fama-miller working paper.

Gilchrist, S., and E. Zakrajsek (2012): "Credit Spreads and Business Cycle Fluctuations," *American Economic Review*, 102(4), 1692–1720.

Golosnoy, V., B. Gribisch, and R. Liesenfeld (2012): "The Conditional Autoregressive Wishart Model for Multivariate Stock Market Volatility," *Journal of Econometrics*, 167(1), 211–223.

Gourieroux, C., J. Jasiak, and R. Sufana (2009): "The Wishart Autoregressive Process of Multivariate stochastic Volatility," *Journal of Econometrics*, 150(2), 167–181.

Gourio, F. (2012): "Disaster Risk and Business Cycles," *American Economic Review*, 102(6), 2734–66.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78(3), 402–17.

Greenwood, J., Z. Hercowitz, and P. Krusell (1997): "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, 87(3), 342–62.

Guerrieri, L., and M. Iacoviello (2013): "Collateral constraints and macroeconomic asymmetries," Discussion paper.

Guerron-Quintana, P. A. (2010): "What you match does matter: the effects of data on DSGE estimation," *Journal of Applied Econometrics*, 25(5), 774–804.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007): "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, 54(8), 2291 – 2304.

Hamilton, J. D. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357–84.

Hansen, P. R., Z. Huang, and H. H. Shek (2012): "Realized GARCH: a Joint Model for Returns and Realized Measures of Volatility," *Journal of Applied Econometrics*, 27(6), 877–906.

Hansen, P. R., A. Lunde, and J. M. Nason (2011): "The Model Confidence Set," *Econometrica*, 79(2), 453–497.

Harvey, D., S. Leybourne, and P. Newbold (1997): "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13(2), 281 – 291.

Hautsch, N., and Y. Ou (2012): "Analyzing Interest Rate Risk: Stochastic Volatility in the Term Structure of Government Bond Yields," *Journal of Banking & Finance*, 36(11), 2988 – 3007.

Hautsch, N., and F. Yang (2012): "Bayesian Inference in a Stochastic Volatility Nelson–Siegel Model," *Computational Statistics & Data Analysis*, 56(11), 3774 – 3792.

Herbst, E., and F. Schorfheide (2014): "Bayesian Inference for DSGE Models," Discussion paper.

Jacquier, E., N. G. Polson, and P. E. Rossi (1994): "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business & Economic Statistics*, 12(4), 371–89.

Jaimovich, N., and S. Rebelo (2009): "Can News about the Future Drive the Business Cycle?," *American Economic Review*, 99(4), 1097–1118.

Jermann, U., and V. Quadrini (2012): "Macroeconomic Effects of Financial Shocks," *American Economic Review*, 102(1), 238–71.

Jin, X., and J. M. Maheu (2013): "Modeling Realized Covariances and Returns," *Journal of Financial Econometrics*, 11(2), 335–369.

Jungbacker, B., S. J. Koopman, and M. van der Wel (2013): "Smooth Dynamic Factor Analysis with Application to the US Term Structure of Interest Rates," *Journal of Applied Econometrics*.

Jurado, K., S. C. Ludvigson, and S. Ng (2013): "Measuring Uncertainty," NBER Working Papers 19456, National Bureau of Economic Research, Inc.

Justiniano, A., G. Primiceri, and A. Tambalotti (2011a): "Investment Shocks and the Relative Price of Investment," *Review of Economic Dynamics*, 14(1), 101–121.

——— (2011b): "Investment Shocks and the Relative Price of Investment," *Review of Economic Dynamics*, 14(1), 101–121.

Justiniano, A., and G. E. Primiceri (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98(3), 604–41.

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2010): "Investment shocks and business cycles," *Journal of Monetary Economics*, 57(2), 132–145.

Karapanagiotidis, P. (2012): "Improving Bayesian VAR Density Forecasts through Autoregressive Wishart Stochastic Volatility," MPRA Paper 38885, University Library of Munich, Germany.

Khan, H., and J. Tsoukalas (2012): "The Quantitative Importance of News Shocks in Estimated DSGE Models," *Journal of Money, Credit and Banking*, 44(8), 1535–1561.

Kim, C.-J. (1994): "Dynamic linear models with Markov-switching," *Journal of Econometrics*, 60(1-2), 1–22.

Kim, J., and F. J. Ruge-Murcia (2009): "How much inflation is necessary to grease the wheels?," *Journal of Monetary Economics, Elsevier*, 56(3), 365–377.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, 65(3), 361–393.

Koenker, R. W., and J. Bassett, Gilbert (1978): "Regression Quantiles," *Econometrica, Econometric Society*, 46(1), 33–50.

Koopman, S. J., M. I. P. Mallee, and M. Van der Wel (2010): "Analyzing the Term Structure of Interest Rates Using the Dynamic Nelson–Siegel Model With Time–Varying Parameters," *Journal of Business & Economic Statistics*, 28(3), 329–343.

Koopman, S. J., and E. H. Uspensky (2002): "The Stochastic Volatility in Mean Model: Empirical Evidence from International Stock Markets," *Journal of Applied Econometrics*, 17(6), 667–689.

Litterman, R. B. (1979): "Techniques of forecasting using vector autoregressions," Discussion paper.

Liu, Z., D. F. Waggoner, and T. Zha (2011): "Sources of macroeconomic fluctuations: A regimeswitching DSGE approach," *Quantitative Economics*, 2(2), 251–301.

Maheu, J. M., and T. H. McCurdy (2011): "Do High-Frequency Measures of Volatility Improve Forecasts of Return Distributions?," *Journal of Econometrics*, 160(1), 69 – 76.

Morley, J., and J. Piger (2012): "The Asymmetric Business Cycle," *The Review of Economics and Statistics*, 94(1), 208–221.

Neftci, S. N. (1984): "Are Economic Time Series Asymmetric over the Business Cycle?," *Journal of Political Economy, University of Chicago Press*, 92(2), 307–28.

Philipov, A., and M. E. Glickman (2006): "Multivariate Stochastic Volatility via Wishart Processes," *Journal of Business & Economic Statistics*, 24(3), 313–328.

Primiceri, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72(3), 821–852.

Rinnergschwentner, W., G. Tappeiner, and J. Walde (2011): "Multivariate Stochastic Volatility via Wishart Processes - A Continuation," Working Papers 2011-19, Faculty of Economics and Statistics, University of Innsbruck.

Rotemberg, J. J. (1982): "Sticky Prices in the United States," *Journal of Political Economy*, 90(6), 1187–1211.

Ruge-Murcia, F. (2012a): "Estimating nonlinear DSGE models by the simulated method of moments: With an application to business cycles," *Journal of Economic Dynamics and Control*, 36(6), 914–938.

Ruge-Murcia, F. J. (2012b): "Skewness Risk and Bond Prices," Discussion paper.

Saijo, H. (2013): "The Uncertainty Multiplier and Business Cycles," UTokyo Price Project Working Paper Series 016, University of Tokyo, Graduate School of Economics.

Schmitt-Grohe, S., and M. Uribe (2012): "What's News in Business Cycles," *Econometrica*, 80(6), 2733–2764.

Schorfheide, F. (2005): "Learning and Monetary Policy Shifts," *Review of Economic Dynamics*, 8(2), 392–419.

Shephard, N., and K. Sheppard (2010): "Realising the Future: Forecasting with High-

Frequency-Based Volatility (HEAVY) Models," *Journal of Applied Econometrics*, 25(2), 197–231.

Sims, C. A., and T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy?," *American Economic Review*, 96(1), 54–81.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

Takahashi, M., Y. Omori, and T. Watanabe (2009): "Estimating Stochastic Volatility Models Using Daily Returns and Realized Volatility Simultaneously," *Computational Statistics & Data Analysis*, 53(6), 2404 – 2426.

Tauchen, G. (1986): "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20(2), 177–181.

Uhlig, H. (1997): "Bayesian Vector Autoregressions With Stochastic Volatility," *Econometrica*, pp. 59–73.

van Dijk, D., S. J. Koopman, M. van der Wel, and J. H. Wright (2013): "Forecasting Interest Rates with Shifting Endpoints," Tinbergen Institute Discussion Papers 12-076/4, Tinbergen Institute.